

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: May 7, 2004, 11:55:46 ; Search time 50 Seconds

(without alignments)  
5616.224 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890  
Sequence: 1 MFPAAPAPKRWLPFLILLILL.....CNKALKPDAKPCBSQLCPL 890

Scoring table: Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	3.4	183	11	Q99JPI mus musculus
2	13	1.5	2165	5	Q19791 caenorhabditis
3	12	1.3	833	11	Q8K384 mus musculus
4	12	1.3	845	11	Q8BNJ2 mus musculus
5	11	1.2	129	12	Q8VBA0 white spot
6	11	1.2	166	11	Q80Z90 mus musculus
7	11	1.2	197	11	Q80Z91 mus musculus
8	11	1.2	230	11	Q80Z92 mus musculus
9	11	1.2	263	11	Q60819 mus musculus
10	11	1.2	406	10	Q80L88 oryza sativa
11	11	1.2	558	10	Q7XU06 oryza sativa
12	10	1.1	83	3	Q870N2 neurospora
13	10	1.1	143	4	Q15412 homo sapien
14	10	1.1	165	4	Q726V1 homo sapien
15	10	1.1	192	6	Q95N24 equus caball
16	10	1.1	203	11	Q8CH80 cavia porce

17	10	1.1	245	4	Q72502 homo sapien
18	10	1.1	271	6	Q95KQ0 pan paniscu
19	10	1.1	274	6	Q9GKD9 bos taurus
20	10	1.1	278	4	Q9BTO9 homo sapien
21	10	1.1	292	6	Q9GKE0 bos taurus
22	10	1.1	322	7	Q8KMW1 bos taurus
23	10	1.1	322	7	Q9GIZ5 homo sapien
24	10	1.1	322	10	Q9SHN5 mus musculus
25	10	1.1	324	7	Q95IB2 homo sapien
26	10	1.1	324	7	Q95HN6 homo sapien
27	10	1.1	347	5	Q81209 plasmodium
28	10	1.1	413	10	Q84PU6 oryza sativa
29	10	1.1	414	6	Q8HZR0 canis fami
30	10	1.1	439	11	Q8K466 mus musculus
31	10	1.1	439	11	Q8BLH5 mus musculus
32	10	1.1	464	13	Q9DGD1 gallus gall
33	10	1.1	467	4	Q8NEX4 homo sapien
34	10	1.1	467	4	Q9BTY2 homo sapien
35	10	1.1	467	4	Q726Y2 homo sapien
36	10	1.1	481	5	Q18122 caenorhabdi
37	10	1.1	492	11	Q7IPW7 mus musculus
38	10	1.1	521	11	Q08795 mus musculus
39	10	1.1	528	11	Q921X2 mus musculus
40	10	1.1	548	4	Q9BR08 homo sapien
41	10	1.1	559	11	Q8BRU4 mus musculus
42	10	1.1	606	6	Q97554 cryocolagus
43	10	1.1	633	6	Q8HZR1 canis fami
44	10	1.1	759	6	Q8HZM8 equus caball
45	10	1.1	769	11	Q912V2 rattus norv

## ALIGNMENTS

RESULT 1  
ID Q99JPI PREDIMINARY; PRT; 183 AA.  
AC Q99JPI  
DT 01-JUN-2001 (TREMURel, 17, Created)  
DT 01-JUN-2001 (TREMURel, 17, Last sequence update)  
DT 01-OCT-2003 (TREMURel, 25, Last annotation update)  
DE Hypothetical protein (fragment).  
GN ADAMTS8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strauberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC005780; AA05780.1; -  
DR MGD; MGI:1353468; Adamts8.  
DR InterPro; IPR000884; TSP1.  
DR SMART; SM00209; TSP1; 1.  
DR PROSITE; PS50092; TSP1; 1.  
KW Hypothetical protein.  
FT NON\_TER  
SQ SEQUENCE 183 AA; 19960 MW; 5C15HC3E09B3FB3 CRC64;  
Query Match 3.4%; Score 30; DB 11; Length 183;  
Best Local Similarity 100.0%; Pred. No. 7.6e-22;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 741 GQYLLNGNLAIISAIEDIIVKGTILKXSGS 770  
Db 34 GQYLLNGNLAIISAIEDIIVKGTILKXSGS 63  
RESULT 2  
ID Q19791 PRELIMINARY; PRT; 2165 AA.  
AC Q19791; Q27524;

```

DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE F25H8.3 protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajdedy S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copesey T., Cooper C., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten N., Latreille P.,
RA Lighthouse J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaiden N., Smith A., Sonhammer E., Straden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstein L., Wilkinson-Sprat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RT Nature 368:32-38 (1994).
RN EMBL; Z69361; CA93288.1; -.
DR EMBL; Z69360; CA93288.1; JOINED.
DR EMBL; Z69360; CA93287.1; -.
DR EMBL; Z69361; CA93287.1; JOINED.
DR PIR; T21371; T21371.
DR HSBP; P15167; IDTH.
DR MEROPS; M12.135; -.
DR WormPep; F25H8.3; CE05729.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 14.
DR SMART; SM00209; TSP1; 18.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 15.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 2165 AA; 244397 MW; FCC3DA8AA9C4888 CRC64;
Query Match 1.5%; Score 13; DB 5; Length 2165;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 536 WGECSTRTGGGVQ 548
DB 612 WGECSTRTGGGVQ 624
RESULT 3
Q8K384 PRELIMINARY; PRT; 833 AA.
AC Q8K384;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to a disintegrin-like and metalloprotease (reprolysin type)
DE with thrombospondin type 1 motif, 4.
GN ADAMTS4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027773; AAH27773.1; -.
DR MGI; MGI:133949; Adamts4.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007223; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 14.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 18.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 15.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KM Integrin; Protease; Metalloprotease.
SQ SEQUENCE 833 AA; 90097 MW; 4CEC83DFEC3AA619 CRC64;
Query Match 1.3%; Score 12; DB 11; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 539 CSRTGGGVQPS 550
DB 526 CSRTGGGVQPS 539
RESULT 4
Q8BNJ2 PRELIMINARY; PRT; 845 AA.
AC Q8BNJ2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease.
GN ADAMTS4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK083534; BAC38944.1; -.
DR MGI; MGI:133949; Adamts4.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 14.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.

```

DR PROSITE; PSS0215; ADAM\_MERPO; 1.  
 DR PROSITE; PSS0092; TSPI; 1.  
 DR PROSITE; PSS0142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 845 AA; 91339 MM; B496C3190D1A9225 CRC64;

Query Match 1.3%; Score 12; DB 11; Length 845;  
 Best Local Similarity 100.0%; Pred. No. 0.01;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGVOFS 550  
 DB 540 CSRTCGGVOFS 551

RESULT 5  
 Q8VBA0 PRELIMINARY; PRT; 129 AA.

AC Q8VBA0; 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE W5V080 (W5SV137).  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.  
 OX NCBI\_TaxID=92652;

RN [1] SEQUENCE FROM N.A.  
 RX MEDLINE=2154831; PubMed=11689662;  
 RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
 RT "Complete genome sequence of the shrimp white spot bacilliform virus";  
 RL J. Virol. 75:11811-11820(2001).

RN [2] SEQUENCE FROM N.A.  
 RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [3] SEQUENCE FROM N.A.  
 RP STRAIN=Taiwan;  
 RX MEDLINE=20517548; PubMed=11062040;  
 RA Tsai W.F., Yu H.T., Tzeng H.F., Lwu J.H., Chou C.M., Huang C.J.,  
 Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
 RT "Identification and characterization of a shrimp white spot syndrome virus (WSSV) gene that encodes a novel chimeric polypeptide of cellular-type thymidine kinase and thymidylate kinase";  
 RL Virology 277:100-110(2000).

RN [4] SEQUENCE FROM N.A.  
 RP STRAIN=Taiwan;  
 RX MEDLINE=21844071; PubMed=11853398;  
 RA Chen L.L., Lei J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
 Lo C.F., Kou G.H.;  
 RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white spot syndrome virus and characterization of the motif important for targeting VP35 to the nuclei of transfected insect cells";  
 RL Virology 293:44-53(2002).

RN [5] SEQUENCE FROM N.A.  
 RP STRAIN=Taiwan;  
 RA Lo C.-F., Kou G.-H.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF432093; AAL3084.1;  
 DR EMBL; AF440570; AAL89005.1;  
 SQ SEQUENCE 129 AA; 14606 MM; 2B436884C1948180 CRC64;

Query Match 1.2%; Score 11; DB 12; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 0.02;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 FLULLLLPL 23  
 DB 114 FLULLLLPL 124

RESULT 6  
 Q80Z90 PRELIMINARY; PRT; 166 AA.

DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Interleukin 15 receptor alpha chain isoform 1C.  
 GN IL15RA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6;  
 RA Toomey J.A., Gays F., Foster D., Brooks C.G.;  
 RT "Cytokine requirements for the growth and development of mouse NK cells in vitro";  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY219717; AA062312.1;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam; PF00084; sush1.1.  
 DR SMART; SM00032; CCP; 1.  
 KW Receptor.  
 SQ SEQUENCE 166 AA; 18122 MM; 94DC3A60B7CD883B CRC64;

Query Match 1.2%; Score 11; DB 11; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLPL 24  
 DB 18 LLLLLLLPL 28

RESULT 7  
 Q80Z91 PRELIMINARY; PRT; 197 AA.

DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Interleukin 15 receptor alpha chain isoform 1B.  
 GN IL15RA.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6;  
 RA Toomey J.A., Gays F., Foster D., Brooks C.G.;  
 RT "Cytokine requirements for the growth and development of mouse NK cells in vitro";  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY219716; AA062311.1;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam; PF00084; sush1.1.  
 DR SMART; SM00032; CCP; 1.  
 KW Receptor.  
 SQ SEQUENCE 197 AA; 21366 MM; 1D16A75C518B986B CRC64;

Query Match 1.2%; Score 11; DB 11; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 0.03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLPL 24  
 DB 18 LLLLLLLPL 28

## RESULT 8

Q80292 PRELIMINARY; PRT; 230 AA.

AC Q80292; 080292;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Interleukin 15 receptor alpha chain isoform 1A.  
 GN IL15RA.  
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCB1\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6;

RA Toomey J.A., Gays F., Foster D., Brooks C.G.;

RT "Cytokine requirements for the growth and development of mouse NK

RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY219715; A062310.1; -

DR GO; GO:0004872; F:receptor activity; IEA.

DR InterPro; IPR000436; Sush1\_SCR\_CCP.

DR Pfam; PF00084; Sush1; 1.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

Query Match 1.2%; Score 11; DB 11; Length 263;  
 Best Local Similarity 100.0%; Pred. No. 0.039;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLLL 24  
 DB 18 LLLLLLLLLL 28

## RESULT 10

Q8S0L8 PRELIMINARY; PRT; 406 AA.

AC Q8S0L8; 08S0L8;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Putative pectinacetylesterase.  
 GN B1078G07.15.

OS Oryza sativa (japonica cultivar-group).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Eriacridaceae; Oryzaceae; Oryza.  
 CX NCB1\_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA Sasaki T., Matsunoto T., Yamamoto K.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC

RL clone:B1078G07."

DR Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP003407; BAB90194.1; -

DR Gramene; O8S0L8; -

DR InterPro; IPR004963; Pec\_acetylase.

DR Pfam; PF03283; PAB; 1.4579 MW; C47926770E80A5CB CRC64;

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.

DR SMART; SM00032; CCP; 1.

DR Receptor.



Query Match 1.2%; Score 11; DB 10; Length 558;  
 Best Local Similarity 100.0%; Pred. No. 0.076;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 15 LLLLLLLLLPLA 25  
 |||||  
 DB 7 LLLLLLLLLPLA 17

## RESULT 12

0870N2 PRELIMINARY; PRT; 83 AA.

AC Q870N2; 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein 49D12.210.  
 GN 49D12.210.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

OX NCB1\_TaxID=5141;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Schulte U., Aligh V., Hobeisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.,  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BX25540; CAD79696.1; -  
 KW Hypothetical protein.

SQ SEQUENCE 83 AA; 9293 MW; 8B6B07546FCE6A CRC64;

Query Match 1.1%; Score 10; DB 3; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 LLLLLLLLLP 23  
 |||||  
 DB 18 LLLLLLLLLP 27

## RESULT 13

015412 PRELIMINARY; PRT; 143 AA.

AC 015412; 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE CPG4A.  
 GN CPG4A.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCB1\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain cortex;  
 RX MEDLINE=97369492; PubMed=9225980;  
 RA Margolis R.L., Abraham M.R., Gatchell S.B., Li S.H., Kidwai A.S.,  
 RA Breschel T.S., Stine O.C., Gallahan C., McInnis M.G., Rose C.A.,  
 RT "CDNs with long CAG trinucleotide repeats from human brain."  
 DR HUM. Genet. 100:114-122(1997).  
 DR EMBL; U080744; AAB91442.1; -  
 SQ SEQUENCE 143 AA; 15959 MW; 0F2119BP4B33A1E6 CRC64;

Query Match 1.1%; Score 10; DB 4; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 LLLLLLLLLP 23  
 |||||  
 DB 17 LLLLLLLLLP 26

## RESULT 14

Q726V1 PRELIMINARY; PRT; 165 AA.

AC Q726V1; 01-OCT-2003 (TREMBlrel. 25, Created)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE DJ20N2.5.2 (Novel protein (MG1314) similar to fucosidase, alpha-L-1,  
 DE tissue, variant 2).  
 GN DJ20N2.5.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCB1\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Kay M.,  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL031320; CAD92485.1; -  
 SQ SEQUENCE 165 AA; 19835 MW; 0427DB86CD215820 CRC64;

Query Match 1.1%; Score 10; DB 4; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.27;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 LLLLLLLLLP 23  
 |||||  
 DB 13 LLLLLLLLLP 22

## RESULT 15

Q95N24 PRELIMINARY; PRT; 192 AA.

AC Q95N24; 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Aggrecaenase-1 (Fragment).

OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCB1\_TaxID=9796;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Flannery C.R., Little C.B.,  
 RT "Expression and activity of equine aggrecaenases".  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF368321; AAK53425.1; -

DR MEROPS; M1.221; -  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR PROSITE; PS0215; ADAM\_MEROP; 1.  
 DR PROSITE; PS0142; ZINC\_PROTEASE; 1.  
 FT NON TER 1 1  
 FT NON TER 192 192  
 SQ SEQUENCE 192 AA; 20670 MW; 9013B0E19FCE8C56 CRC64;

Query Match 1.1%; Score 10; DB 6; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 495 PMADGTPCPG 504  
 |||||  
 DB 157 PMADGTPCPG 166

Fri May 7 12:18:01 2004

us-09-989-687-4.rpt

Page 6

Search completed: May 7, 2004, 12:00:09  
Job time : 52 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: May 7, 2004, 11:43:25 ; Search time 18 seconds

(without alignments)  
2574.580 Million cell updates/sec

Title: US-09-989-687-4  
4853  
Sequence: 1 MFPAAPAPRMFLPILLILL.....CNKATKPEDAKPCESQICPL 890

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	4853	100.0	890 1	ATP8_HUMAN
2	3918.5	80.7	905 1	ATP8_MOUSE
3	2344	48.3	968 1	ATP8_MOUSE
4	2343.5	48.3	967 1	ATP8_MOUSE
5	2328.5	48.0	967 1	ATP8_MOUSE
6	2138.5	44.1	950 1	ATP8_MOUSE
7	1842.5	38.0	837 1	ATP8_MOUSE
8	1768.5	36.5	930 1	ATP8_MOUSE
9	1734.5	35.7	930 1	ATP8_MOUSE
10	1720.5	35.5	1906 1	ATP8_MOUSE
11	1699.5	35.0	1935 1	ATP8_MOUSE
12	1644.5	33.9	630 1	ATP8_MOUSE
13	1633	33.6	1911 1	ATP8_MOUSE
14	1284.5	26.5	562 1	ATP8_MOUSE
15	1200	24.7	997 1	ATP8_MOUSE
16	1189	24.5	1593 1	ATP8_MOUSE
17	1124.5	23.2	1081 1	ATP8_MOUSE
18	1108	22.8	1077 1	ATP8_MOUSE
19	1108	22.8	1211 1	ATP8_MOUSE
20	1102	22.7	1205 1	ATP8_MOUSE
21	1072.5	22.1	1213 1	ATP8_MOUSE
22	1070	22.0	1224 1	ATP8_MOUSE
23	1049	21.6	1205 1	ATP8_MOUSE
24	1016.5	20.9	1223 1	ATP8_MOUSE
25	1004	20.7	1095 1	ATP8_MOUSE
26	989	20.4	1210 1	ATP8_MOUSE
27	979	20.2	860 1	ATP8_MOUSE
28	971	20.0	1207 1	ATP8_MOUSE
29	615.5	12.7	245 1	ATP8_MOUSE
30	538	11.1	207 1	ATP8_MOUSE
31	498	10.3	377 1	ATP8_MOUSE
32	488.5	10.1	525 1	ATP8_MOUSE
33	428.5	8.8	824 1	ATP8_MOUSE

34	344.5	7.1	920 1	ADP2_MOUSE
35	341.5	7.0	776 1	ADP2_MOUSE
36	336.5	6.9	813 1	ADP2_MOUSE
37	332.5	6.9	956 1	ADP2_MOUSE
38	330.5	6.8	793 1	ADP2_MOUSE
39	325	6.7	819 1	ADP2_MOUSE
40	323	6.7	775 1	ADP2_MOUSE
41	318	6.6	857 1	ADP2_MOUSE
42	317.5	6.5	754 1	ADP2_MOUSE
43	314	6.5	864 1	ADP2_MOUSE
44	311	6.4	909 1	ADP2_MOUSE
45	309.5	6.4	826 1	ADP2_MOUSE

## ALIGNMENTS

RESULT 1  
ID ATP8\_HUMAN  
AC Q9UP79; Q9NZS0;  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2) (METH-8).  
GN ADAMTS8 OR METH2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=93367466; PubMed=10438512;  
RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S., Lombardo M., Iruela-Arispe M.L.;  
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity.";  
RL J. Biol. Chem. 274:23349-23357(1999).  
RN [2]  
RP SEQUENCE OF 195-440 FROM N.A.  
RX MEDLINE=20079168; PubMed=10610729;  
RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;  
RT "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on mouse chromosome 9 and human chromosome 11.";  
RL Genomics 62:312-315(1999).  
RN [3]  
RP FUNCTION: Has anti-angiogenic properties.  
CC - FUNCTION: Binds 1 zinc ion per subunit (By similarity).  
CC - SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).  
CC - TISSUE SPECIFICITY: Highly expressed in adult and fetal lung, lower expression in brain, placenta, heart, stomach and fetal brain and kidney.  
CC - DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.  
CC - PTM: The precursor is cleaved by a furin endopeptidase (By similarity).  
CC - SIMILARITY: Belongs to peptidase family M12B.  
CC - SIMILARITY: Contains 1 disintegrin-like domain.  
CC - SIMILARITY: Contains 2 TSP type-1 domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL: AF060153; AAD48081.1; -  
CC EMBL: AF175283; AAF25806.1; -

DR HSSP; P34179; 11AG.  
 DR MEROPS; M12.226; -.  
 DR Genew; HGNC:224; ADAMTS8.  
 DR MIM; 605175; -.  
 DR GO; GO:0005178; F: integrin binding; TAS.  
 DR GO; GO:0008237; F: metalloproteinase activity; TAS.  
 DR GO; GO:0007345; F: embryogenesis and morphogenesis; TAS.  
 DR GO; GO:0008285; P: negative regulation of cell proliferation; TAS.  
 DR InterPro; IPR001762; D: Disintegrin.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; TSP1; 2.  
 DR PRINTS; PRO1705; TSP1REPEAT.  
 DR SMART; SM00209; TSP1; 2.  
 DR PROSITE; PS00215; ADAM\_MEROPS; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS00214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS00092; TSP1; 2.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR HydroLase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KM Repeat; Extracellular matrix; Heparin-binding.  
 FM SIGNAL 1 27  
 FM PROPEP 28 214  
 FM CHAIN 25 890  
 FM DOMAIN 215 430  
 FM DOMAIN 439 526  
 FM DOMAIN 527 582  
 FM DOMAIN 584 690  
 FM DOMAIN 691 832  
 FM DOMAIN 834 889  
 FM DOMAIN 889 202  
 FM METAL 364 364  
 FM ACT SITE 365 365  
 FM METAL 368 368  
 FM METAL 374 374  
 FM CARBOHYD 345 345  
 FM CARBOHYD 401 401  
 FM CARBOHYD 466 466  
 FM CARBOHYD 491 491  
 FM CARBOHYD 600 600  
 FM CONFLICT 195 195  
 FM CONFLICT 413 440  
 SQ SEQUENCE 890 AA; 96671 MW; 57D0EB03D5739D3 CRC64;  
 Query Match 100.0%; Score 4853; DB 1; Length 890;  
 Best Local Similarity 100.0%; Pred. No. 2e-293;  
 Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPPAPAPRMLPRLILLILLILLPLARAGAPAPPAQASGLVYPTPLPSAGELMLHISA 60  
 DB 1 MFPPAPAPRMLPRLILLILLILLPLARAGAPAPPAQASGLVYPTPLPSAGELMLHISA 60  
 QY 61 FGKFFVRLAPDDSLAPEFKIRLGSGGATGEGRLGCGFSGTVNGEPESLAAVSIC 120  
 DB 61 FGKFFVRLAPDDSLAPEFKIRLGSGGATGEGRLGCGFSGTVNGEPESLAAVSIC 120  
 QY 61 FGKFFVRLAPDDSLAPEFKIRLGSGGATGEGRLGCGFSGTVNGEPESLAAVSIC 120  
 DB 61 FGKFFVRLAPDDSLAPEFKIRLGSGGATGEGRLGCGFSGTVNGEPESLAAVSIC 120  
 QY 121 RGLSGSLDGEETITPOGAGSLAOPHRLQWGPAGAPLPFRGPEWEVTEGEGORER 180  
 DB 121 RGLSGSLDGEETITPOGAGSLAOPHRLQWGPAGAPLPFRGPEWEVTEGEGORER 180  
 QY 121 RGLSGSLDGEETITPOGAGSLAOPHRLQWGPAGAPLPFRGPEWEVTEGEGORER 180  
 DB 121 RGLSGSLDGEETITPOGAGSLAOPHRLQWGPAGAPLPFRGPEWEVTEGEGORER 180  
 QY 181 GDHODESEESQSEEAAGSEPPPLGATSTRKFVSEARFVETLLVADASAAAFYADL 240  
 DB 181 GDHODESEESQSEEAAGSEPPPLGATSTRKFVSEARFVETLLVADASAAAFYADL 240  
 QY 181 GDHODESEESQSEEAAGSEPPPLGATSTRKFVSEARFVETLLVADASAAAFYADL 240  
 DB 181 GDHODESEESQSEEAAGSEPPPLGATSTRKFVSEARFVETLLVADASAAAFYADL 240  
 QY 241 QNHLLTMSVAARIYKPKSTKNSINLMVYVLLVEDEKGPBPVSDNGGILLTNFCWQOR 300  
 DB 241 QNHLLTMSVAARIYKPKSTKNSINLMVYVLLVEDEKGPBPVSDNGGILLTNFCWQOR 300

QY 301 FNPDSRHPHEDYDTAILLITRONFCGQEGCPTLGVADIGTICDPKSCSVIIEGLOAH 360  
 DB 301 FNPDSRHPHEDYDTAILLITRONFCGQEGCPTLGVADIGTICDPKSCSVIIEGLOAH 360  
 QY 361 TLAEHGLVLSNPHDSDKCTFLFPGMKGHVMAFLFHLNQTLPMSCSAMYITELLDG 420  
 DB 361 TLAEHGLVLSNPHDSDKCTFLFPGMKGHVMAFLFHLNQTLPMSCSAMYITELLDG 420  
 QY 421 GHSDCLLDAPGALPLPTGLPGRVAYLOLDQCRQIFGPPDRHCENPISAQDVCAQLMCHT 480  
 DB 421 GHSDCLLDAPGALPLPTGLPGRVAYLOLDQCRQIFGPPDRHCENPISAQDVCAQLMCHT 480  
 QY 481 DGAEPLCHTKNSLPMADSTPCGPHLCSGSCLPBEEVERPKFVDDGMAPWGWGBCS 540  
 DB 481 DGAEPLCHTKNSLPMADSTPCGPHLCSGSCLPBEEVERPKFVDDGMAPWGWGBCS 540  
 QY 541 RTGGGVQFSHRECKDPFQNGARYCLGRPAKYQSCHTBEGCPDQKSFREOCCERYANN 600  
 DB 541 RTGGGVQFSHRECKDPFQNGARYCLGRPAKYQSCHTBEGCPDQKSFREOCCERYANN 600  
 QY 601 YTMDSNLLQWPVKYAGVSPRCKLFCRARSEFKVFEKVIDGTLCGPELTALCVRG 660  
 DB 601 YTMDSNLLQWPVKYAGVSPRCKLFCRARSEFKVFEKVIDGTLCGPELTALCVRG 660  
 QY 661 QCVKAGCDHVSPKRLDKGVCGGKNSCRKVSGLTPTVGYNDITIPAGATNIDVK 720  
 DB 661 QCVKAGCDHVSPKRLDKGVCGGKNSCRKVSGLTPTVGYNDITIPAGATNIDVK 720  
 QY 721 QRHREPVQNDGYVLLKTDGQYLNLGALSAISQDILVYGTILKYSATLIERLOSF 780  
 DB 721 QRHREPVQNDGYVLLKTDGQYLNLGALSAISQDILVYGTILKYSATLIERLOSF 780  
 QY 781 RPLREPLVQLITVPEVFPKRYKTFPVNDVDFSMQSKERATNIIQPLHAAWYLG 840  
 DB 781 RPLREPLVQLITVPEVFPKRYKTFPVNDVDFSMQSKERATNIIQPLHAAWYLG 840  
 QY 841 DMEGCSSTGAGMORRTVECRDPSGQASATCNKLPKPADPCSOQCEL 890  
 DB 841 DMEGCSSTGAGMORRTVECRDPSGQASATCNKLPKPADPCSOQCEL 890

RESULT 2  
 ID\_MOUSE STANDARD; PRT; 905 AA.  
 AC P87110;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).  
 GN ADAMTS8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20079168; PubMed=10610729;  
 RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;  
 RT "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on RT mouse chromosome 9 and human chromosome 11.";  
 RL Genomics 62:312-315(1999).  
 CC -1- FUNCTION: Has anti-angiogenic properties (By similarity).  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed specifically in adult lung and heart and low expression during mouse development.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix. (By similarity).  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 2 TSP type-1 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR HSMB; AFI75282; AAF5805.1; -.  
 DR HSMB; P34179; 11AC.  
 DR MEROPS; M12.226; -.  
 DR WGD; MGI1353468; Adamts8.  
 DR Interpro; IPR001762; Disintegrin.  
 DR Interpro; IPR006025; Pept\_M Zn BS.  
 DR Interpro; IPR001590; Peptidase\_M12B.  
 DR Interpro; IPR000884; TSP1.  
 DR Interpro; IPR000885; TSP1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; Tsp\_1; 2.  
 DR PRINTS; PRO1705; TSP1REPEAT.  
 DR SMART; SMO0209; TSP1; 2.  
 DR PROSITE; PSS0215; ADAM\_MEROPS; 1.  
 DR PROSITE; PSS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PSS0214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PSS0092; TSP1; 2.  
 DR PROSITE; PSS0142; ZINC\_PROTEASE; 1.  
 DR PROSITE; PSS0142; ZINC\_PROTEASE; 1.  
 DR HydroLase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KM Repeat; Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 28  
 FT PROPEP 29 228  
 FT CHAIN 229 905  
 FT DOMAIN 229 452  
 FT DOMAIN 453 541  
 FT DOMAIN 542 597  
 FT DOMAIN 599 705  
 FT DOMAIN 706 847  
 FT DOMAIN 848 904  
 FT METAL 378 379  
 FT ACT\_SITE 379 379  
 FT METAL 382 382  
 FT METAL 388 388  
 FT CARBOHYD 415 415  
 FT CARBOHYD 480 480  
 FT CARBOHYD 506 506  
 FT CARBOHYD 615 615  
 SQ SEQUENCE 905 AA; 96879 MW; 124D4132B3A0CAE CRC64;  
 Query Match 80.7%; Score 3918.5; DB 1; Length 905;  
 Best Local Similarity 79.2%; Pred. No. 1.7e-235;  
 Matches 727; Conservative 74; Mismatches 99; Indels 15; Gaps 4;  
 QY 1 MFPAAPAPRLPFLILLILL-LLEPLARCAPAPPAAGGSELVPRPLPGSGELALHLS 59  
 DB 1 MURDPTTGTGMPPLILLILLPPPLVCGAPAGGTAAQSELVPRPLPGSASSELAFHLS 60  
 QY 60 AFSGKGVRLRAPDSEFLAEFEKIERLGSGRATGCRGGRGFFSGTVNGEPESLAAVSL 119  
 DB 61 AFGQGVRLRAPDSEFLAEFEKIERLGSGSAAAGSPGARGFSGTVNGEPESLAAVSC 120  
 QY 120 CRGLSGSFLLDGESEFTIQPGAGSIAQPHRLQRMKPAAR-----PLRGPWM 168  
 DB 121 VAWMSGSFLLAGSEFTIQPGAGSIAQPHRLQRMKPGRRDPGLAAAEVPLDQLEW 180  
 QY 169 EVELTGSGORERGDHDESEESQE--EBAEGASEPPPLGATSRTRFVSEARFVETLL 226  
 DB 181 EVMGNGGQGERGSDNEDRKQDEKGLKETEDSRKXVPPFGSKTRKRVSEARFVETLL 240  
 QY 227 VADASMAAFYGADLOHNLITLMSVAAIRYKHSIKNSILNLYVKKVILVEDKMGSPVSDN 286

DB 241 VADASMAAFYGTDIQNHILITMSMAARIYKHSIRNSVNLVYVKVILVEKMGSPVSDN 300  
 QY 287 GGLTLRFNFCMORREPNPSDRPHRYDTAILLTRNFCGSGGLCDITIGVADIGTIDPDK 346  
 DB 301 GGLTLRFNFCMORREPNPSDRPHRYDTAILLTRNFCGSGGLCDITIGVADIGTIDPDK 360  
 QY 347 SCSTVEDEGLQAATLHLELGHVLSMHPDSEKPTPLFGPKGKHVWAPLFVHLNQLTPW 406  
 DB 361 SCSTVAKBEGLOAAVTLHLELGHVLSMHPDSEKPTPLFGPKGKHVWAPLFVHLNQLTPW 420  
 QY 407 SPCAMVYTELLEDGSHDCLIDAPGALPLPTGLPGEMALYQIDQRCQITFGPDFRCEN 466  
 DB 421 SPCAVYTELLEDGSHDCLIDAPGALPLPTGLPGEMALYQIDQRCQITFGPDFRCEN 480  
 QY 467 TSAQDVCAQIMC-HTGAEPICHTKNGSLPMADGTPCGPHGLCSGSCLPREEVERPKV 525  
 DB 481 TSVEDICVQLCARBDEDEPTCHTKNGSLPMADGTPCGPHGLCLDGSCLVEDENPKAV 540  
 QY 526 VDGMAPWPGWEGESRFGCGGVSPSHRECKDPEFONGRVCYIGBRATYQSCNTECEPPG 585  
 DB 541 VDGMPGWRPFGWQSRFGCGGVSPSHRECKDPEFONGRVCYIGBRATYQSCNTECEPPG 600  
 QY 586 KSFEQCEKYNAYNTDMGNILLQWPKYAGVSPDRCKLFCRARGSEKVEEAYID 645  
 DB 601 KSFEQCEKYNAYNTDMGNILLQWPKYAGVSPDRCKLFCRARGSEKVEEAYID 660  
 QY 646 GTLCGPETLACVAGCCVKAQCDHVNSPKLDCGVCGGKNGSCRYVSGSLPTNYGVN 705  
 DB 661 GTLCGPETLACVAGCCVKAQCDHVNSPKLDCGVCGGKNGSCRYVSGSLPTNYGVN 720  
 QY 706 DIVTIPAGATNIDVQKSHRPGVNDGNVLLKTAQGYLLNGNLASAIEDILVKGITL 765  
 DB 721 DIVTIPAGATNIDVQKSHRPGVNDGNVLLKTAQGYLLNGNLASAIEDILVKGITL 780  
 QY 766 KYSSGSIATLELQSFRLPEPLTYQLTLVPEGEVPPPKYKTFPVPNDVDSMOSKERAT 825  
 DB 781 KYSSGSIATLELQSFRLPEPLTYQLTLVPEGEVPPPKYKTFPVPNDVDSMOSKERAT 840  
 QY 826 TNIQPLHMQWVGDMSGSGSTCGAGMORRTVECRPSQASATCKNAKLPEDAKPES 885  
 DB 841 TNIQPLHMQWVGDMSGSGSTCGAGMORRTVECRPSQASATCKNAKLPEDAKPES 900  
 QY 886 QLCPL 890  
 DB 901 QLCPL 905  
 RESULT 3  
 ATSL\_MOUSE STANDARD; PRT; 968 AA.  
 ID ATSL\_MOUSE  
 AC P97857; OS4768;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motif 1) (ADAM-TS 1) (ADAM-TS1).  
 GN ADAMTS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=128/SVJ;  
 RX MEDLINE=96110583; PubMed=9441751;  
 RA Kuno K., Liza H., Ohno S., Matsushima K.;  
 RT "The exon/intron organization and chromosomal mapping of the mouse  
 RL ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";   
 RV Genomics 46:466-471(1997).  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97150761; PubMed=8995297;  
 RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,



```

QY 124 SGGFLDGEFTTPOGAGSLAQPHRQWGPAGAPPLPRGEMEVETGEG-----ORQE 179
Db 153 RGAFFYLQGEFFFIQPP-----GVATERLAPNAYEESSARPOPHILRRR 198
QY 180 RGP-----HOEDSEESQDEEAGSEPPPLGA-----TSRTKAPVSEA 219
Db 199 RGGGAKCGCWDEDELPTSDRPESONTRNQPVRDPDQAGKPSGSIKKKRFVSSP 258
QY 220 RFEETLLVADAAFAFGADLOHILTLMSVAARIKHSINKSNILMAYVYKLIYEDDEM 279
Db 259 RYEVMTLVADQSMADPHSGSLKAYLLITLSVAARFKHSINKSNISLVVYKLIYEBEK 318
QY 280 GPEVSNNGITLTFNFCNWRFPNOPSDRPHENYTAILLTRONFCGSEGLCDTLGVADIG 339
Db 319 GPEVTNNAALITLNFNCWOKHNSPSDRDBEHYDTAILFTRODLGSH--TCDTLGMADV 377
QY 340 TIGDPKSGSVLEDEGLQAAHTLAHLGHVLSMHPHDSKRCRRLFGPMKHHVMAPLFVH 399
Db 378 TVCDPERSGSVLEDDGLQAAFTTAHLGHVFNPHDAGHCSLNGVTGDSHLSMSS 437
QY 400 LNCITLWSPSCSANYITELL.DGSHGCDLLDAPGALPLPTLPGRNALYOLDQCRQITGP 459
Db 438 LHSQGWSPSCSAYMTVTSFLDNHGECLMDXPNPIKLPSDLG--TLVDANNQCFITGE 495
QY 460 DPHRCNTSAQDVCAQLNC--HTDGAAPLCHTKNGSLPMADGTPCGPRLHLSGSLPEBE 518
Db 496 ESKHCPD--AASTCTTLMCTGTSGLLVQYQYH--FPMADGTSCEGKVCVSGKCVNTXD 551
QY 519 VERPKFVNDGMAFPMGECSCRTCGGYQVSHRECKDEPEONGRGVCLGRKATYQSCHT 578
Db 552 MGHFATPVNAGSNCPMGMDCSRTCGGYQVMEEDNDVFNKNGKCYEGKRVYRSCNI 611
QY 579 EBCPP--DGSGFRBQCEKTNAYNTYDMGN--LLQWPKYAGVSPDRCKLFCRARGSE 635
Db 612 EBCPDNNGKTFREEOCEANHEFSKSF--GNEPTEVETPKYAGVSPDRCKLFCRARGSE 670
QY 636 FKVEFAKVIDGTLGCEETLALCVRGQCVKAGDHVDSRKLDKCVGCKGNSCRKYSK 695
Db 671 FVYLQKAVVDGTPCSBDSVSCVQGGCVAGCDRIITDSKKRDKCGVCGNSSTCKRMSG 730
QY 696 SLPTNYGVNDIVTIPAGATNIDVKKSHPGVQNDGNYLALXTADGOYLLNGLNLSAIE 755
Db 731 IYVSTRPGHDIYTIIPAGATNIEVKHNRGSRNNGSFLAIPADGTIILNGLNLSLIE 790
QY 756 QDLIVAGTILTKYSGSIALTERLQSPRLPEPLTVQLTYPGVEFPKXKTYTFVNDVDF 815
Db 791 QDLTYRGVTLARYSGSAALETRISFPLKEPLTIQVLMV--GHALRPKIKTYTFM----- 843
QY 816 SMOSKERATVNIICPLHAQWVLGDMWSECSSTCGAGMORRTVECRDPSGASATCNKAL 875
Db 844 ----KKTESFPAIPIPF--SEWVIEBWGCSKTCGSGWRVYQGRDINGHNASCAKEV 897
QY 876 KPEDAKPCESQUCP 889
Db 898 KPASTRPCADLPCP 911

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Brain;
RA Liu X., Yu Y., Yin T., Johnstone E.W., Stephenson D.T., Clemens J.A.,
RA Little S.P.;
RT "Induction of a disintegrin and metalloprotease with the
RT thrombospondin type I motif (ADAMTS).";
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
RM [2]
RP SEQUENCE OF 18-967 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Liver;
RX MEDLINE=20304099; PubMed=10847486;
RA Diamantis I., Luetth M., Hoevel M., Reichen J.;
RT "Cloning of the rat ADAMTS-1 gene and its down regulation in
RT endothelial cells in cirrhotic rats.";
RL Liver 20:165-172(2000).
CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. Has angiogenic inhibitor activity (by
CC similarity). Active metalloprotease, which may be associated with
CC various inflammatory processes as well as development of cancer
CC cachexia. May play a critical role in follicular rupture (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-Leu-1684
CC site, within the chondroitin sulfate attachment domain.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -1- INDUCTION: Down-regulated in endothelial cells derived from
CC cirrhotic liver.
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF149118; AAD34012.1; -
DR EMBL; AF304446; AAG29823.1; -
DR MEROPS; M12.222; -
DR InterPro; IPRO06586; ADAM cysteine.
DR InterPro; IPRO01762; Disintegrin.
DR InterPro; IPRO06025; Pept_M12B_BS.
DR InterPro; IPRO01590; Peptidase_M12B.
DR InterPro; IPRO02870; Peptidase_M12B_N.
DR InterPro; IPRO00884; TSP1.
DR InterPro; IPRO08085; TSP_1.
DR Pfam; PF01562; Pept_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS00215; ADAM_MEROP; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS50093; TSP1; 3.
KW Hydrolyase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 54
FT PROPEP 55 252
FT CHAIN 253 967
FT DOMAIN 253 475
FT DOMAIN 476 558
FT DOMAIN 559 614
FT TSP TYPE-1 1.

```

FT DOMAIN 616 724 CYS-RICH.  
 FT DOMAIN 725 857 SPACER.  
 FT DOMAIN 854 910 TSP TYPE-1.2.  
 FT DOMAIN 911 967 TSP TYPE-1.3.  
 FT DOMAIN 194 198 POLY-ARG.  
 FT SITE 205 205 CYSTEINE SWITCH (POTENTIAL).  
 FT METAL 401 401 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 402 402 BY SIMILARITY.  
 FT METAL 405 405 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 411 411 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 21 21 I -> V (IN REF. 2).  
 FT CONFLICT 26 31 KFRSSQ -> RSRGSL (IN REF. 2).  
 FT CONFLICT 49 49 V -> A (IN REF. 2).  
 FT CONFLICT 72 72 R -> P (IN REF. 2).  
 FT CONFLICT 79 79 L -> TR (IN REF. 2).  
 FT CONFLICT 249 249 R -> G (IN REF. 2).  
 FT CONFLICT 262 265 TMV -> NLK (IN REF. 2).  
 FT CONFLICT 607 607 S -> F (IN REF. 2).  
 FT CONFLICT 936 936 L -> V (IN REF. 2).  
 FT CONFLICT 962 962 I -> T (IN REF. 2).  
 SQ SEQUENCE 967 AA; 105705 MW; F93C864F6DCDB4CF CRC64;  
 Query Match 48.3%; Score 2343.5; DB 1; Length 967;  
 Best Local Similarity 49.5%; Pred. No. 8e-138;  
 Matches 451; Conservative 143; Mismatches 243; Indels 75; Gaps 19;

DB 613 CPDNNKTFREDECEANERSKSE - GNEPTWETPKYAGVSPDRCKLTCEAKGICGYF 671  
 QY 638 VFEAKVIDGTLGCEETLAIQVRCQCVKAGCDHVVDSPKLDKCGVCGKNGSKRYSGSL 697  
 DB 672 VLQKRVVDGTFCSDSISVCCQCVKAGCDRIIDSKKFKDGCVCNGNSTCKKISGTV 721  
 QY 698 TPTNYGNDIVTTAGATNIDYKORSHPGVQNDGNATLAKTADQVYLNNLNLISAIED 757  
 DB 732 TSTPGYNDIVTTAGATNIEVKHNPGRSGNNSSFLAIPADQFTYLLNGFTLSTLEOD 791  
 QY 758 IIVGTLIKYGSATLTERLQSPRLPEPLTVQLTVPGEVFPKVKYTFEPNDVDSM 817  
 DB 792 LTVGATVRYGSSAALERIRSFSLKEPLIQVLAV - GHALRKIKYTFE - 842  
 QY 818 QSSKERATTNIQPLNADQVYLGWSGCCSTCGAGMORRYECRDPGGAATCNKALKP 877  
 DB 843 -KKTEBFNAPIPF--SEWVIEEMGCSKTCGSGWQRPVECDINGHPASBCAKEVKP 898  
 QY 878 EDAPCESQLCP 889  
 DB 899 ASTRPCLDLPCEP 910  
 RESULT 5  
 ID ATSL HUMAN STANDARD; PRT; 967 AA.  
 AC Q9UJH8; Q9NSJ8; Q9P2K0; Q9UH83; Q9UP80;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).  
 GN ADAMTS1 OR METH1 OR KIAA1346.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;  
 RT "Cloning, characterization and mapping on human chromosome 21 of the  
 RT orthologue of murine Adams-1.";  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RA Tissue=Heart;  
 RX MEDLINE=99367466; PubMed=10438512;  
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,  
 RA Lombard M., Inoué-Arispe M.L.;  
 RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new  
 RT family of proteins with angio-inhibitory activity.";  
 RL J. Biol. Chem. 274:23349-23357(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Tissue=Endothelial cells;  
 RX MEDLINE=20247184; PubMed=10785405;  
 RA Gienke J., Schmitt A.O., Pilarsky C., Hinemann B., Weiss B.,  
 RA Rosenthal A., Thierach K.H.;  
 RT "Differential gene expression by endothelial cells in distinct  
 RT angiogenic states.";  
 RL Eur. J. Biochem. 267:2820-2830(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Tissue=Brain;  
 RX MEDLINE=20181126; PubMed=10718198;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XVI.  
 RT The complete sequences of 150 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 7:65-73(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953;





```

QY 582 PP-DGKFRROCKKXVANYTDX-DGNLLOWPKYAGVSPRRCKLFCRARGSEFKYE 639
DB 614 PDNNGKTFREOCEAHNEFSKASGSPAYEMIPKTAGVSPKRCULICQAKIGFFFL 673
QY 640 EAYVDTLGGPELIALCVRGQCYKAGCDHVVDSFRKLDKGVCGGKNSCRKVSGLTP 699
DB 674 QPKVDTGSPDSTSCVCGQCYKAGCDRIIDSKKFKDCGVCGGNSCKKISGSVTS 733
QY 700 TNYGNDVITPAPATNDVYKORSHPGVONDVIAKTDAGGYLLNGNIAIAIDOLI 759
DB 734 AKPEYHIIITPTATNIEVAKORNSRNGSFLAKADGYIILNGDITSLIDOLIM 793
QY 760 VKGTLIKSGSIATLERLQSRPLPEPLVOLTVPGSEVPPPKYKTFPVPNDVDSKOS 819
DB 794 YKGVVLRSGSSALERTSRSPSLKEPLTIGVLTIV-GNALRPKIKYTFV----- 842
QY 820 SKKATNTIITPLHAGVNLGDNSECSSTCGAGMQRRTVCRDPSGQASATCKALKPED 879
DB 843 KKKKESFNAL-PTFSA-WVLEWGECSKSCELGMQRVLVECRDINGPASECAKEVPAS 900
QY 880 AKPCEQQLCP 889
DB 901 TRPCADHPCP 910

RESULT 6
AT15_HUMAN STANDARD; PRT; 950 AA.
AC O8TE58;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-15 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
  with thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15).
GN ADAMTS15.
NS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
  Gal S., Oaya A.J., Iamazares M., Garabaya C., Quesada V.,
  Lopez-Otin C.;
  "Cloning, expression analysis, and structural characterization of
  seven novel human ADAMTS, a family of metalloproteinases with
  distinct integrin and thrombospondin-1 domains.";
  Gene 283:49-62(2002).
RL Gene 283:49-62(2002).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
  matrix (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in fetal liver and kidney, but not
  in any of the adult tissues examined.
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
  for a tight interaction with the extracellular matrix (By
  similarity).
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
  similarity).
CC -1- SIMILARITY: Belongs to peptidase family M2B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation-
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC EMBL: AJ315733; CAC86014.1; -.
CC DR MEROPS: M12.025; -.
CC DR Genew; HGNC:16305; ADAMTS15.

```

```

DR MIM; 607509; -.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001818; Pept M10A_M12B.
DR InterPro; IPR006025; Pept M Zn BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01562; Pept_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; TSP_1; 3.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS50215; ADAM_MEROPS; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;
  Repeat; Extracellular matrix.
FT SIGNAL 1
FT PROPEP 18
FT CHAIN 212
FT DOMAIN 213
FT DOMAIN 213
FT DOMAIN 427
FT DOMAIN 428
FT DOMAIN 516
FT DOMAIN 571
FT DOMAIN 572
FT DOMAIN 701
FT DOMAIN 838
FT DOMAIN 839
FT DOMAIN 896
FT SITE 174
FT METAL 361
FT ACT_SITE 362
FT METAL 365
FT METAL 371
FT CARBOHYD 141
FT CARBOHYD 591
FT CARBOHYD 623
FT CARBOHYD 679
SQ SEQUENCE 950 AA; 103286 MW; 5DFBE18285CCCC3B CRC64;

Query Match 44.1%; Score 2138; DB 1; Length 950;
Best Local Similarity 45.3%; Pred No. 4e-125;
Matches 426; Conservative 150; Mismatches 255; Indels 109; Gaps 24;

14 LILLILLILLPLARGAPAPPAAGQAS-ELVPTSL-----PGSAGE-LALH 57
1 MILLIGILLAF-----GRVAGSEPEREVVPIRLDPINGRRYRYPEDSDGLIFQ 56
58 LSAFGKGFVRLAPDPSFLAPFKTERLGSGRA-TGGERGLRGCFSGTVNGEPESIAA 116
57 ITAFQEDFYHLTLPAGFLAPAFSTEHGVPLQGLGSSDLRCFPGDVAAEPDSFRA 116
117 VSLCRGLSGSFLLDSEFTICP-QGAGGSLAQ-----PRLQR----- 153
117 VSLCGSLRGAFGVRGAVYISPLPNVASAPAAQNSGAHLQRPVGGSGDPTSRGCV 176
154 ---WGRAGKRL-PROPEWVETGSGQORERDHDQEDSESEGEAEASAPPPPLIAT 209
177 ASGMNPAIRLADDPYKFR--RAGFQESRSR-----RRS 207
210 SRTKRFVSEARFEVETLLVADASMAAFYGADLGNHLLTMSVARIRYKPSIKNSINLVV 269
208 GRKRFVSLPRVETLLVADSEWVFHGADLHHVLLTLATARLYRHSILNPNIVVV 267
270 KVLIVDEKQGEVSDNGELTIRNFCNWRPNRQSDRPHRYDRAILLTRONFGQSEGL 329
268 KYLLLDROSGEPTVGNALLTRNFCWQKLNKXSDRPHRYDRAILLTRDDLCGAT-T 326
330 CDPLGVADIGTICDPKNSGVTEDEGLQAHTLAELGVLSPHSDSKPCTRLFGPMCK 389
327 CDPLGVADIGTICDPKNSGVTEDEGLQAHTLAELGVLSPHSDSKPCTRLFGPMCK 386
390 HHVMAPLFVHLNQLTFLWSPCSAMYLTELLDGGHGDCLLDAPGAALPLPTGLFGRAALYQL 449

```



```

FT DOMAIN 213 436 METALLOPROTEASE.
FT DOMAIN 437 519 DISINTEGRIN-LIKE.
FT DOMAIN 520 575 TSP TYPE-1.
FT DOMAIN 577 685 CYS-RICH.
FT DOMAIN 686 837 SPACER.
FT DOMAIN 837 252 POLY-ALA.
FT SITE 194 194 CYSTEINE SWITCH (POTENTIAL).
FT METAL 361 361 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 362 362 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 365 365 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 371 371 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 77 77 A -> T (IN REF. 1).
FT CONFLICT 626 626 R -> Q (IN REF. 3).
FT CONFLICT 682 682 G -> R (IN REF. 3).
SQ SEQUENCE 837 AA; 90224 MW; 5DF9C9AC137DF41F CRC64;

Query Match 38.0%; Score 1842.5; DB 1; Length 837;
Best Local Similarity 45.0%; Pred. No. 7,7e-107;
Matches 375; Conservative 130; Mismatches 244; Indels 85; Gaps 24;

5 PAAP-RWLPULLILLI-LIPLRGAAPAPAGQSELVPRRLPGSA-----GELA 55
27 PIYPLSLWVLLILLILSLPSARLASPLP-----REBEIVFPKRLNGSVLPQSGAPARLL 82
56 LHLSPGKGFVRLAPDPSFLAPEFKTERLGSGGRATGGERGLRGCFSGSTVNGSEESIA 115
83 CRLOAFETLLLEBDSQVQVEGLTVYQGAPELLGAE--PGTYLTGTINGDESEVA 140
116 AVSLCRG-LSGSFLDDEEFTIQ-----PQAGGSLAQPHRLQWGPAGAPPLPGPEW 168
141 SLHMDGALGLVLYRGAEFLHQLLEGSTPNSAGGPGA--HILRRKSPASG-----QGPMC 194
169 EYVGEQQRQERGHQDSESESEBEAGSEPPPLGATS-----RTKRFVEAREFET 224
195 NVKX-----PLGSPSPRRPRARFALSFRVET 222
225 LVLVADMAAFYAGADLQNHILTLMSVARIYKHSIKNSINLMVVKYLIVEDSKWPEVS 284
223 LVVADDMKAAFHAGLKRVLTVWAAAKAFKHSINPVLVTRVLVILGSEEBPQVQ 282
285 DNGSLTLRNFQNRQRRNPSDRPHRYDAILLTRNFGQSGELCDTIGVADIGTICP 344
283 PSAQTRKSPFAMWGRGINTPDSPPDHPDAILFTRDCLG--VSTCTTLMADVGVCD 341
345 NKSGSVLEDEGLQAHLAHLAHLGVLSPHDSKPCRLFGPMG-KHIVAPLFLVHLNQT 403
342 ARSCAIVDDDLQSAFLTAHLGHVFMMLHNSKPCISLNGPLSTSHWAPVMAHVDE 401
404 LPWSPGSAHYTLLELDGHCCLLDADQALPLFTGLPGMALYQLDQCRQITFGPDFR 463
402 EPMWPCSAKRLITDLNNGYGHCLLDKPEAPLHLVTPPGXD--VDARQQLTFFGPDSR 469
464 CPNPSAOCVCAQLWC--HTDGAELCHTNGSLPMADGTFCGPHGLCSBSCPLPEEVER 521
460 CPQLPBP--CALWCSGLNG--HAMCQTKHS--PMAQGTGCPAQACMGKRCILMDQLD 514
522 PKPVVDGMAWGWGEGSRTCCGGGVQFSHECEDPEPQNGRCCLRRAKQSCPTREC 561
515 FNIDQAGMGWGPWGDCSRTCCGGGVQFSRHCTRPVRNGKXCCEBRKRRFRSCNEDC 574
582 PP-DKSPFROQCEKYNAHYTMDGNL---LQWVPKAGVSPRDKCTLCRAARGSEEF 637
575 PTGSLTLTREFQCAAYN--HRTDLFKSPGPMADVPRYTGAAPDQCKLTCQARALGY 632
638 VFEAKVDTGLGPEPTALCYRGQCVAKGCHVVDSPKLDKCVGCGKNSCKKVSGL 657
633 VLEPRVVDGTCPSPSSSVQVQGRICRRIIGSKKKKFKCMVCCGDSGSGSKGGSF 692
698 TPTNGYNDIYITIPGATINIVKORSHGVQNDGNVYALAKTADQYCLNGVLAISAEOD 757
693 RKFRIGYNNVYITIPGATHILVROQNGN--HRSIYLAALKLPDSVYALNGEYTLMPSP 750

```

```

QY 758 ILVKGCTI-LKXSGSIATLERLQSRPLPEPTVQGLTVPGEVPEPKYKTFEVP 810
DB 751 VLVPAVSLRYSKATPAASETSGHGFLAQPLTGLV--VAGNPQTRILRYSEFVP 803

RESULT 8
ATSS_MOUSE
ID ATSS_MOUSE STANDARD; PRT; 930 AA.
AC Q9R001
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2) (ADMP-2) (Implantin).
DE ADAMTS5.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99395124; PubMed=10464288;
RA Hukakinen T.U., Hirohata S., Seidlin M.F., Apre S.S.;
RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteinases."
RL J. Biol. Chem. 274:25555-25563 (1999).
CC - FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be involved in its turnover. May play an important role in the destruction of aggrecan in arthritic diseases. May play a role in proteolytic processing mostly during the peri-implantation period.
CC - CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-Ala-393 site.
CC - COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC - SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
CC - DEVELOPMENTAL STAGE: Expressed specifically in the peri-implantation period in embryo and trophoblast and at low or undetectable level thereafter.
CC - DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.
CC - PTM: The precursor is cleaved by a furin endopeptidase (By similarity).
CC - SIMILARITY: Belongs to peptidase family M12B.
CC - SIMILARITY: Contains 1 disintegrin-like domain.
CC - SIMILARITY: Contains 2 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC EMBL: AF140673; AAD56356.1; -
CC MGD: MGI:1346321; Adamts5.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR006025; Pept_M12B.
DR InterPro: IPR001590; Peptidase_M12B.
DR InterPro: IPR002870; Peptidase_M12B_N.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR008085; TSP 1.
DR Pfam: PF01562; Pcp_M12B_propep. 1.
DR Pfam: PF01421; Repolysin; 1.
DR Pfam: PF00090; TSP1_1_2.
DR PRINTS: PR01705; TSP1RBPAT.
DR SMART: SM00209; TSP1_2.
DR PROSITE: PS00215; ADAM_MPRO; 1.
DR PROSITE: PS00092; TSP1_2.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.

```

KM Hydrolase; Metalloprotease; zinc; signal; Glycoprotein; Zymogen;  
 KM Repeat; Extracellular matrix.  
 FT SIGNAL 1 21  
 FT PROPEP 22 261  
 FT CHAIN 262 930  
 FT DOMAIN 262 484  
 FT DOMAIN 262 484  
 FT DOMAIN 566 485  
 FT DOMAIN 567 622  
 FT DOMAIN 624 731  
 FT DOMAIN 732 874  
 FT DOMAIN 732 874  
 FT DOMAIN 875 929  
 FT DOMAIN 929 929  
 FT SITE 257 261  
 FT SITE 209 209  
 FT METAL 410 410  
 FT ACT SITE 411 411  
 FT METAL 414 414  
 FT METAL 420 420  
 FT CARBOHYD 498 498  
 FT CARBOHYD 728 728  
 FT CARBOHYD 802 802  
 FT CARBOHYD 807 807  
 SQ SEQUENCE 930 AA; 101780 MM; 84DE84B26170D4DC CRC64;  
 Query Match 36.5%; Score 1769.5; DB 1; Length 930;  
 Best Local Similarity 40.3%; Pred. No. 2,9e-102;  
 Matches 381; Conservative 124; Mismatches 355; Indels 85; Gaps 19;  
 10 WIPFLLILLL---LPLFARGA---RPAAGQASELVLP-----44  
 5 WASLLILLLLASCLSLAADSFAAPADKTRQPOAAAAAEPDQGEETREGRHLQ 64  
 45 -----TRLPSSAGELALHLSAFGKGVFLRLAPDPSFLAFEFKTERLGSGR 90  
 65 LAGRRSGGLVNIIDQLVSGGKGVIVYAGRRRLDLDERDT-VGAAGSIYVAGGLS 123  
 91 ATGERGLRGCFSGTGVNGEPESLAVALICRLGSSFLIDGEBFTLOQAGGSLACPHR 150  
 124 ASSGRG--HCFRGTVDGSPRSLVFDLCGLDGFVAKARYLKLPL-LRGSVAEYER 180  
 151 LORWGPAGAR-----PLPRGMEVEVEGEQORERGDHQDESEESQ-----192  
 181 I--YDGGSRILHYNNRBSFEALPPRASCTPASPQSPQSPVHSRRRLAPQL 238  
 193 -EEABGASEPPPLGATSRTPFVSEARFVETLLVADASMAAFYADLQNHILTVMSVA 251  
 239 LDHSAFSPSGNAGPQTWRRRRRSISRAROVELLIVADSSMAWYGRGLQHYLLTLASIA 298  
 252 ARIYHPSIKNSINIMVYVLLVEDEKMGPEVSDNGGLTLNFCMORRFPQSPRPEH 311  
 299 NLYSHASIEHRIKLAIVKVVUTDKDTSLEVSKAATTLTKFCMKQGNHQLGDHREH 358  
 312 VDTAILLRONFCGEGELCDTLGVADIGTICDPNKSQSVIIEDEGLQAHTLAHELGHVLS 371  
 359 YDAALIFREDICGHHN-CDTLGMADVGTICSPESCAVIBDDGHAFAVHAHEIGHILG 417  
 372 MEHDSKPETRLFGPMGKHVMAPLFVHNOTLPPSPASAWLTIELIDGSGDCLTADG 431  
 418 LSHDSKCEENFGTTEDEKRLMSILTSIDASKPSKTSATITFFLDGNGCLDLDR 477  
 432 AALPLPTGLPGMALYOLDQCGROIFGDFRHCPTMSAQDYCAOLMCH-TDGAELPCHTK 490  
 478 KQILPEELPQO--TYDATQCNTLTFGEYSVCP--GMDVCARLMCAVAGQGVCLTK 532  
 491 NQSLMAGTGPCGPHLCSGSCLEBEVERKPVVDGMAFWMGWGCSTTCGGGQFOS 550  
 533 --KLPAVBSTPCGKGRVCLQCKVCKTKKYYSTSSHNWGSWGWGCGSSCGGQVFA 590  
 551 HEECDPEPONGRCYCLGRAYQSCHTBECPPDQSGFREDOCEKYNAVYNTDMG--NL 608  
 591 YAHGNPAPRNSGRCTGKRAIYRSCSVTPCPNKSFRHNGCEAKNGYQ--SDAKGYKTF 649  
 609 LOWPKYAGVSPRDRCKLFCRARGSEBRVPEAKYIDGTLGPEETLAICTVAGQCVKAGCD 668

Db 650 VEWPKTAGVLPADVCLTRPRAGTGYVVFSPVDTGTECRYSNSVCRAGCVRTGCD 709  
 Qy 669 HVVDSPEKLDKCGVCGKGNSSCRVSGSLTPTYGVNDIYTPAGATNIDVKORSHGVQ 728  
 Db 710 GIISKLYQYKCGVCGGDDNSCTKIKIGTFNKKSQGYDVRIEAGATHIVRQPKADQ 769  
 Qy 729 NDGNYLALTKADQYLLNGLAISALEQDILVYGTILKYSGLATLERL--QSFRPLPEP 786  
 Db 770 RFPAYLALKKKTBEYLNGKXMTSTSETIIDINGTVWNSGWSHRDPLHGMYSATKEI 829  
 Qy 787 LTVQLTVPEGEVPPK---VKYTFEVPNDVDFSGQSSKERATTNIIQPLHAQVLDMS 843  
 Db 830 LIQGLATD---PTALGVRYSFVPKTKTKVNSVISHGSKVGHSTQLOMVGPM 885  
 Qy 844 EGSSTCAGVQRTVEGRDPSGQASATCNKALPEADKPCESQ 888  
 Db 886 ACSRTCDTGMHRTVOCODGNRLKAGCCLISQPSAFKQCLTKKC 930  
 RESULT 9  
 ACSS\_HUMAN  
 ID ACSS\_HUMAN STANDARD; PRT; 930 AA.  
 AC Q9UNAO; Q9UNP2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2) (ADMP-2) (ADAM-TS 11).  
 GN ADAMTS OR ADMP2 OR ADAMTS11.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=9367476; PubMed=10438522;  
 RA Abbaszade I., Liu R.-O., Yang F., Rosenfeld S.A., Ross O.H., Link J.R., Ellis D.M., Tortorella M.D., Pratta M.A., Hollis J.M., Wynn R., Duke J.L., George H.J., Hillman M.C., Jr., Murphy K., Wsawall B.H., Copeland R.A., Decicco C.P., Bruckner R., Nagase H., Ito Y., Newton R.C., Magolda R.V., Trzaskos J.M., Hollis G.F., Arner E.C., Burn T.C.;  
 RA "Cloning and characterization of ADAMTS11, an aggrecanase from the ADAMTS family.";  
 RL J. Biol. Chem. 274:23443-23450 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953;  
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenblatt A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordstrek G., Hornischer K., Brandt P., Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H., Ramser J., Beck A., Klages S., Hennig S., Rasseilmann L., Dagand E., Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lebrun H., Reinhardt R., Yaspo M.-L.;  
 RA "The DNA sequence of human chromosome 21.";  
 RL Nature 405:311-319 (2000).  
 RN [3]  
 RP SEQUENCE OF 413-930 FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=99395124; PubMed=10464288;  
 RA Hurlstaine T.L., Hirohata S., Seldin M.F., Apté S.S.;  
 RT "ADAM-TS, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteases.";  
 RL J. Biol. Chem. 274:25555-25563 (1999).

CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover. May play an important role in the  
 CC destruction of aggrecan in arthritic diseases. May play a role in  
 CC proteolytic processing mostly during the post-implantation period.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-Ala-393  
 CC site.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (by similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed at low level, primarily in placenta  
 CC but also in other tissues, such as heart and brain, and also  
 CC cervix, uterus, bladder, esophagus, rib cartilage,  
 CC chondroblastoma, fibrous tissue and joint capsule from an  
 CC arthritic patient.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (by  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF142039; AAD9577.1; -  
 DR EMBL; AP001698; BAA95504.1; -  
 DR EMBL; AP001697; BAA95503.1; -  
 DR EMBL; AF141293; AAF02493.1; -  
 DR HSPB; O9PM35; 1BD.  
 DR MEROPS; M12.225; -  
 DR Genew; HGNC:221; ADAMTSL.  
 DR MIM; 605007; -  
 DR GO; GO:0005578; C:extracellular matrix; TAS.  
 DR GO; GO:0005178; F:integrin binding; TAS.  
 DR GO; GO:0008237; F:metallopeptidase activity; TAS.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M12B.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR008085; TSP 1.  
 DR InterPro; IPR008085; TSP 1.  
 DR Pfam; PF01562; Pept\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reptolysin; 1.  
 DR Pfam; PF00090; TSP\_1\_2; 2.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SMO0209; TSP1; 2.  
 DR PROSITE; PS02015; ADAM\_MEROPS; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE NEG.  
 DR PROSITE; PS02014; DISINTEGRIN\_2; FALSE NEG.  
 DR PROSITE; PS00092; TSP1; 2.  
 DR PROSITE; PS00143; ZINC\_PROTEASE; 1.  
 DR HydroLase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KM Repeat; Extracellular matrix.  
 FT SIGNAL 1 16  
 FT PROPEP 17 261  
 FT CHAIN 262 930  
 FT DOMAIN 262 484  
 FT DOMAIN 485 566  
 FT DOMAIN 567 622  
 FT DOMAIN 624 731  
 FT DOMAIN 732 874  
 FT DOMAIN 875 929  
 FT DOMAIN 930 941  
 FT DOMAIN 942 961  
 FT SITE 257 261  
 FT SITE 209 209  
 FT METAL 410 410  
 ZINC (CATALYTIC) (BY SIMILARITY).

FT ACT SITE 411 411 BY SIMILARITY.  
 FT METAL 414 414 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 420 420 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 728 728 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 802 802 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 807 807 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 138 138 A -> G (IN REF. 2).  
 FT CONFLICT 614 614 P -> L (IN REF. 3).  
 FT CONFLICT 692 692 P -> L (IN REF. 2).  
 SQ SEQUENCE 930 AA; 101715 MW; B64281502F28193B CRC64;  
 Query Match 35.7%; Score 1734.5; DB 1; Length 930;  
 Best Local Similarity 39.8%; Pred. No. 4,3e-100;  
 Matches 378; Conservative 122; Mismatches 348; Indels 101; Gaps 20;  
 15 LLLLLLLPLARGPAPPA---GGC---ASLVVPRRLG----- 49  
 8 LLLCAFLPLAAVGPATPADQXAGPPTAAAAQPRRQGEVQERAPPPHPLAQR 67  
 50 -----SAGELALHSFAGKGVFLRLAPDPSFLAPFKIRIGSGSRATGG 94  
 68 RRSKGVNIDQLXSGGKGVLYVYAGRRLLDLERDGSV-----GIAGVPVAGGG 119  
 95 E---RGIRGCFSGTYNGEPESLAVALSCLRGLSGSLDGEFTIQOGAGSLAOPHR 150  
 120 TSAPVRRSHCFYRGTVASPRSLAVFDLCGLGFFAVKARATLPL-LRGWAEEEK 178  
 151 LQWMPAGARPLPGPEVEVETGSGROE---RGDHOEDSEBSQEEAEGASEPP--- 203  
 179 GRVYDGSARLL-----HYTRRGFSFALPPRASCETPASTPEAHHAHNPGRRA 232  
 204 -----PLGATG-----RTKRVSEARFVETLLVADASAAAFYADLQNHIL 245  
 223 ALASQLDQSALSPAGSGGQPTWRRRRRSISRQVELLVADASAAARLYGRSLCHYLL 292  
 246 TLMSVAARIVYGPISIKYINLMVYKVLVDEKKKGPVSDGGLTLNFCWMOBRFQPS 305  
 293 TLASTANSLYSHASTENIRLAIVKVVVLGDKSLSEVSKAATTLNFCMOHQHQLG 352  
 306 DRHPEHYDTALLTRNFCQEGQLCDTLGVADIGTICDPNKSQVIBDEGLQAHTLAHE 365  
 353 DDHEHYDAALLFRIDLCGHS--CDLGMADVGTICSPERSCAVIDDDGILAAFTVAHE 411  
 366 LGHVLSPHDSKSCCTRLFGPMGKHVMAPIFYHLNQTLLPSPSANYLTLELLDGGHDC 425  
 412 IGHLLGSHDSKSCCTEFTGSTEDEKRLMSLITSIDSKPWSKTSATTEFLDDGHNC 471  
 426 LLDPAGALPLPTGLPGFMALYOLDQCROIFGDFPHCPNTSAQDYCAOLMCH-TDGA 484  
 472 LLDPKQILGPBELPGQ--TYDANQCNLTFGEYSVCP---GMDVCARLMCAVVRQG 526  
 485 PLCHTKNSLPMAGTTPGPGPHLCSBSGLPBEVEVERPKVYVDGMAPMGWMGSCPTG 544  
 527 WCLTK--KLAAVGTGPGKGRICLQGVCKYKTKKYYSTISSHNMWSGWSGQCSHSG 584  
 545 GGVQFSHRECKDPEPONGRYCLGRRAKYQSCHTRECPDGSFRQDCEKXYNNTDM 604  
 585 GGVQAYVHCHNNPAPRNNGRYCTGKRALYRCSLMPCEPNKSRHGECEAKNGYQ--SDA 643  
 605 DG--NLQWPKVAGVSPRDKCLFCARAGSEBKVPEAKYIDGTLGPEFLALCVYGGC 662  
 644 KGVKTFVWVKYAGVLPADVCKLTCAKRGYGVVSPKTYDTECKPYSNSVCVKGKC 703  
 663 VKAGCDHVVDSPRLDKGVCGKSGKSGKRYSGSLFTPNYGYNDIVITIPAGATNIDVQR 722  
 704 VRTGDDGITSGLQYDKGVCGGNSGCKTVGFNNKSKKTYTVLRPEBATHIKRQF 763  
 723 SHPGVQDNVYALKTADQGLNGNTAISAIQDILVKGITLKYSGSIALLEL--QSF 780  
 764 KAKQDRTFAVIALKKNGEYLINGKYMISTSETIILINGVWYSGSHRDEFLHNGY 823  
 781 RPLPEPLTVQLL-TVPGVFPKVKYTFEVENDVDFSQSKERATNIIQPLHAQVVL 839





```

Db      142 TAVFSLGGIMGTGKANDGEYFLEPYLRADGSAHDDHKKPHLYR-----QELKRNSE 195
Qy      166 --PEWEVETGEGORRGDGHQEDSEESQ-----EEBA-----EGASEPPPLGATSRTRYK 215
Db      196 ARBKRCCEVSENMETALPESQSSRTTGDVDLEBAVRLGERSQ-----LHSHKRF 250
Qy      216 VSEARFVETLLVADASMAAFYGDLDONHITLMSVAARYIKPSTKNSINLMVAVYLYE 275
Db      251 LSYRYRYEVVVTADAKVHHGQNLQHYVTLMSIYAAIYKSSIGNLNIIVIKLVVH 310
Qy      276 DEKMGPEVSDNGGLTLRNFGNQRFRNPSDRPEHYDTAILLTRONFGQEGALCDTLGV 335
Db      311 SEGBGVITSNATTLRNFLMGQSQNVPRDAPRSHDRAVILTRDIDGAEKEDTGL 370
Qy      336 ADIGTICDPKSCSVIIEDEGLQAAHTLAEHLGYLSNPHDSCFPCRLRGPMKHHVAP 395
Db      371 AEIGTICDPKSCSVIIEDEGLQAAHTLAEHLGYLSNPHDSCFPCRLRGPMKHHVAP 429
Qy      396 LFNHNLCTLMPSPCSAMVLTLLDGGHGDCLLDAP--GAALPLPTGPGSMALYOLDQQR 454
Db      430 TLNHTSPMTWSACSGKHTTEFLDTGHGECILDKPKRITYDLSPLPG--SVYDGNRQCE 487
Qy      455 QIFGPDPRHCPNTSADVCAQLWC-HTDGAELPCHTKNSLPMADGTPCGPHLSESGC 513
Db      488 LMFPGSQVCPYLKH--CRLWCTSAEYVHKGCRTQH--MPLADGTSQPGWCHRGILC 542
Qy      514 LPEEVEKPEKPVVDGGAHPMGEGCGRTOGGGVQFSHRECHDPEPONGRVCIGSRAX 573
Db      543 VTRDMETRP--VDGEWGPWGPYSSCGRTGGGISTARLCDPEPRNGRVCYVGRMKF 599
Qy      574 QSCHTCECPDGGKSFREQCEKYNANYITMDG--NLLQWPKYAGVSPDRCKLFCRAR 631
Db      600 RSCNTDCEPKGXKDFREKQSDPDKHF-DINGLPVPRVRLPKYSGIAYVRCKLYCRA 658
Qy      632 GRSEPKVEAKVIDTGLCPETLAIQVGGCVAKGCDHVDSFRKIDKGVCGSKNSCR 691
Db      659 GTTSFYOLKDRVADGTPCGTEINDICVQGLCRQAGCHVANSKAKDKCGVCGSDSSQ 718
Qy      692 KVSGLSLPTVNGYNDITYTPAGATNIDVORSHPGVQNDNYALTAKTADQGYLLNGLAI 751
Db      719 TLAIVENSATYGVNIVVXIPGATNIIILHSHSGREDNYIALSDTQGFLLNGLFV 778
Qy      752 SAIBODILVKTILKYSGSIATLTERLQSFRLPEPLTVOLLTPGEVFPKXKITYFVP 810
Db      779 SMAKEINIQGAVEYEGSNNISIERINSTRLAEALVQLCY-GNLYNDVYVSFNIP 837
Qy      811 -----NVDPSMOSKERATT-----NIQPL----- 833
Db      838 EERSNLSFMDYGMQDCTKXCCGLHRRKIACTRKSHAVVSDHNGHLMPLFVTEKCN 897
Qy      834 --HAQWTLGDMSECSSTCGAGMCRRTVEGRDPS--GQA-----SATCNALKPEDAKPC 893
Db      898 MDCLRAHIIIGKSSCCSGCGGYRTLDVHKMKYSVHKGQAVPVQDQCGDQLKPRPREPC 957
Qy      884 ESQICPL 890
Db      958 HGS-CVL 963

```

```

RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORM 3).
RC      TISSUE=Fetal;
RX      MEDLINE=20396138; PubMed=10936055;
RA      Clark M.E., Kellner G.S., Turdeville L.A., Boyer A., Arden K.A.,
RB      Maki R.A.;
RT      "ADAMTS 9", a novel member of the ADAM-TS/Metalloprotein gene
RL      family.
RM      Genomics 67:343-350(2000).
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.
RC      MEDLINE=2251925; PubMed=12514189;
RX      Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,
RA      Evanko S., Wright T.N., Leduc R., Apte S.S.;
RT      "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS
RL      subfamily related to Ctenorhabditis elegans GON-1."
RM      J. Biol. Chem. 278:9503-9513(2003).
RN      [3]
RP      SEQUENCE OF 159-1935 FROM N.A. (ISOFORM 2).
RC      TISSUE=Brain;
RX      MEDLINE=20181126; PubMed=10718198;
RA      Nagase T., Kikuno R., Ishikawa K.-I., Hirogawa M., Ohara O.;
RT      "Prediction of the coding sequences of unidentified human genes. XVI.
RL      The complete sequences of 150 new cDNA clones from brain which code
RM      for large proteins in vitro."
RN      DNA Res. 7:65-73(2000).
CC      -1- FUNCTION: Cleaves the large aggregating proteoglycans, aggrecan
and versican.
CC      -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1838-Glu-Ala-1839
site and versican at the 1428-Glu-Ala-1429 site.
CC      -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC      -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC      -1- ALTERNATIVE PRODUCTS:
Name=1; Synonyms=ADAMTS-9B; Sequence=Displayed;
Name=2; Synonyms=Long;
Name=3; Synonyms=Short;
CC      -1- IsoId=Q9P2N4-1; Sequence=VSP_007548, VSP_007549;
Note=May result from the retention of an intron in the cDNA
leading to a premature stop codon;
Name=3; Synonyms=Short;
CC      -1- TISSUE SPECIFICITY: Highly expressed in all fetal tissues.
Expressed in a number of adult tissues with highest expression in
heart, placenta and skeletal muscle.
CC      -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
for a tight interaction with the extracellular matrix (By
similarity).
CC      -1- DOMAIN: The ancillary domains, including the TSP domain, are
required for specific extracellular localization and for its
versicanase and aggrecanase activities.
CC      -1- PTM: The precursor is cleaved by a furin endopeptidase (By
similarity).
CC      -1- SIMILARITY: Belongs to peptidase family M2B.
CC      -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC      -1- SIMILARITY: Contains 1 GON domain.
CC      -1- SIMILARITY: Contains 15 TSP type-1 domains.
CC      -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC      -1- EMBL: AF261918; AAF89106.1; -
CC      -1- EMBL: AF488803; AA015765.1; -
CC      -1- EMBL: AB037733; BAA92550.1; -
CC      -1- HSSP: P15167; IATL.
CC      -1- MEROPS: M12.021; -
CC      -1- GeneW: HGNC:13202; ADAMTS9.

```





ID AT54 RAT STANDARD; PRT; 630 AA.  
 AC Q95SF7; Q95SP6; Q95SP8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-4 precursor (EC 3.4.24.82) (A disintegrin and metalloproteinase with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1) (Fragment).  
 GN ADAMTS4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mistar; TISSUE=Brain;  
 RA MEDLINE=20415831; PubMed=10961658;  
 RA Satoh K., Suzuki N., Yokota H.;  
 RT "ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin motifs) is transcriptionally induced in beta-amyloid treated rat astrocytes.";  
 RT Neurosci. Lett. 289:177-180 (2000).  
 RL  
 CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be involved in its turnover. May play an important role in the destruction of aggrecan in arthritic diseases. Cleaves aggrecan at the 392-Glu-Ala-393 site.  
 CC -1- CATALYTIC ACTIVITY: Glutamy1 endopeptidase; bonds cleaved include 370-Thr-Glu-Gly-Ala-Arg-Ser-377 in the interglobular domain of mammalian aggrecan.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Only expressed in brain.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.  
 CC -1- PFM: The precursor is cleaved by a furin endopeptidase.  
 CC -1- SIMILARITY: Belongs to peptidase family M2B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL, AB042272; BAB16474.1; -;  
 DR EMBL, AB042271; BAB16473.1; -;  
 DR EMBL, AB042273; BAB16475.1; -;  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR006025; Pept\_Mn\_Bs.  
 DR InterPro: IPR001590; Peptidase\_M2B.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR008085; TSP1.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR Pfam: PF00090; TSP1; 1.  
 DR PRINTS: PR01705; TSP1REPEAT.  
 DR SMART: SM00209; TSP1; 1.  
 DR PROSITE: PS00215; ADAM\_MPRO; 1. FALSE NEG.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE NEG.  
 DR PROSITE: PS00214; DISINTEGRIN\_2; FALSE NEG.  
 DR PROSITE: PS00092; TSP1; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KM Hydroxylase; Metalloprotease; Zinc; Glycoprotein; Zymogen; Extracellular matrix.  
 FT NON\_TER 1  
 FT PROPEP 5  
 FT CHAIN 630 BY SIMILARITY.  
 FT DOMAIN 6 232 ADAMTS-4.  
 FT DOMAIN 233 303 METALLOPROTEASE.  
 FT DOMAIN 313 368 DISINTEGRIN-LIKE.  
 FT TSP TYPE-1.

FT DOMAIN 368 478 CYS-RICH.  
 FT SPACER. 479 630  
 FT DOMAIN 479 630  
 FT POLY-ALA. 480 485  
 FT METAL 154 154  
 FT ACT\_SITE 155 155  
 FT METAL 158 158  
 FT METAL 164 164  
 FT CARBOHYD 96 96  
 FT CARBOHYD 474 474  
 FT SEQUENCE 630 AA; 68384 MW; 63A428753167C7E7 CRC64;  
 SQ  
 Query Match 33.9%; Score 1644.5; DB 1; Length 630;  
 Best Local Similarity 49.9%; Pred. No. 1e-94; Indels 21; Gaps 13;  
 Matches 305; Conservative 107; Mismatches 178;  
 211 RTKRFVSEARFVETLVADASMAAFYGDLOHNTILMSVARIYKPSINSINLVYK 270  
 2 RTKRFASLSRFVETLVADDDKAAAFHAGLKHVLLTVAATAAKAFKPSIRNPVLVYTR 61  
 271 VLTVDKMKGPVPSNGGLTLRNFQNRPNQPSDRRPEHYDTALLTRQNFQGEGLC 330  
 62 LVILSGQEVPPVGSAAQTLSFCTWQKGLNPVDSDDPHDTALLTRQDLG-VSTC 120  
 331 DTLGVADIGTICDPNKSQSVIEDEGLQAHTLAHELGHVLSWPHDSKPCRTLFGP 389  
 121 DALGMAAGVTCDDPARSCAIVEDDLQSAFTPAHELGHVFNMLHDNSKPCANLNGQSSS 180  
 390 HVMAPLFVHLNQLTPNSPCSMYTELIDGHEGCLLDARQALPLPTGLPGRALYQL 449  
 181 RHVMAPVMAVHPPEPWSPCSAFRTIDLDNGYGLLDKPEAPLHLVTPFGKD-YDA 238  
 450 DQCGRIQFQDPDRHCNPNTSAQDVCAQWLC--HTDGAELPLCHTKNGLPWADGTPCGPHL 507  
 239 DQOCUTRFPDSHQPOLPP--CAALMCFGLNG-HMCTKHS--PWADGTPCGPAQA 293  
 508 CSEGSCLPEEVEVERKPVVDGMAWPWGBCSRTGGGVQFSHRECDPEPONGRCL 567  
 294 CMGGRCLHVDQDKDNITDQAGGVPWGDGSRCTGGGVQFSSNDCKRPFERNKGYCE 353  
 568 GRRAYQSGHTEEC-PPGKSPREOQCKRYANVYTDGNNL---LOWPKYAGSPDR 623  
 354 GRTTPRSNTNKPCHGSAITREBOCAALN--HTIDLFKSPGMDVNPRTGTGAPADQ 411  
 624 CLKFCARGSRSEFKFEAKVIDGLCGPETAICVYGQCVAGCDHVVDSPKLDKGVY 683  
 412 CLKTQARALGYVYVLEPRVADGTCSPDSSSVYQGRCHAGCRRILIGSKKKPKCMVC 471  
 684 GKGKNSCRKVGSLPPTNYGVNDITTPAGATNIDVQRSHRGVQNDGNYALAKTXADGY 743  
 472 GNGSSCSKQSGSFKKFRYGSDVVTTPAGRTHILVROQSGSLNS--IYLLKLADSY 529  
 744 LINGNLAIATFODILVKGTI-LKYSGSITATLERLOSFRPELPPTVOLLVPGEVPPK 802  
 530 ALNGEYTLMPSTDVLPAAVSLRYSGRRAAETISGHGFLAQPTLVQL-VAGNPQVNR 588  
 803 VKYTFVENDV 813  
 589 LRYSFVPRPV 599  
 RESULT 13  
 AT20 HUMAN STANDARD; PRT; 1911 AA.  
 ID AT20 HUMAN  
 AC P59510;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-20 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-TS20).  
 GN ADAMTS20.  
 OS Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```

Db 64 QKRSSEALEPMPFRTHYFTAYGQLFQNLTAADASELAAGYBTVHGTPEGAMESDAGP 123
Qy 96 RGLRCFFSCTVNGEPESLAANSICRGSGSFLBGEFTOP-QGAGSGLAQHRLOK 154
Db 124 SDRHCFFRGQNSOEDYKAVVSLCGGLGTGKQNGEYFLEPIKADGN----- 173
Qy 155 GPAGARPLPRGEMEVETEGORERGDHED-----SEESQ----- 192
Db 174 -----EYE-----DGNKHPLIYRDLNNSFLQTLKYCSVSSQIKENSLPHT 217
Qy 193 -----EEBAGASEPPPL---GATSTRTFVSEAFVETLLVADASMAAFGA 238
Db 218 YENMEDNLVWKEKERVIGHTSKNVPLKDERHRSKRRLISYPRYIIMVTAQKAVSAGS 277
Qy 239 DIONHITLMSVAARLYKHPSTKNSINIMVAVLLVEDEKGPEDSDGGLTLNFCNMQ 298
Db 278 NIONYTLTMSIVATLYKDPISIGNLHIVVVLVTHREBEGPVINFGATTLKNFCNMQ 337
Qy 299 REFNPDSRHPHYDTAILLTRONFCGDEGLDYLGVADIGTICDPNKSQVIEDEGLQA 358
Db 338 QTONDLDDVHPSHDITAVLITREDICSSKEKCNITGLSYLTICDPLQSCFINEKGLIS 397
Qy 359 AHTLAHELGVLSMHPDDBSKPTRLFGPMGKHMAAPLVHLNQTLPMSPQSAWLETL 418
Db 398 APTIAHELHTLGVCHDNDPRCKEM-KVTKYHVAAPALSPFMSWMSNSCKRYVIFPL 455
Qy 419 DGHGDCLLDAPGAL-PLPTGLPGSMALYOLDOOCROIFGDPFHCPNTSADVCAGLM 477
Db 456 DRYGSCLLDKDEBEIYNLPSELPG-SRYDNKCELAFGSGQWCPHI---NICMLM 510
Qy 478 C-HTDAEPLCHTNGSLPMADGTPGEPHLCSESGCLPEEVEKPKPVDSGMPW 536
Db 511 CSTETKHKGCFCFQH--VEPADGTGPGMCHRGKCNKETEYTP---VNSEWGPMEPY 565
Qy 537 GECSTCGGAVOPSHCEKDPPEONGRGLGRARYQSCHEBEPDGSFREOCERY 596
Db 566 SSGSRCTGGGIESATRCNRPFRNGVNCVGRMKFRSCNDSCPKQDPFRKQCSDF 625
Qy 597 NAYNTDMDG--NLLQWPKYAGVSPRDCCKFCRARGSEKRVBAKIDITLCGPETL 654
Db 626 NG-KHLDISGIPSNVWMLPRYSGLTKRCKLYCQVAGNIFYLLKDMWEDGTGCTETH 684
Qy 655 AICVRGCVYAGCDHYVDSPRKLDKCGVCGKANSCKRYSGSLTPTNYGNDIVTIPAGA 714
Db 685 DTCVQGCMAAGCDHYLNSAKIDKCGVCGGNSCKITITGVFNSSHVYVNVKIPAGA 744
Qy 715 TNIDVQSRHPVQNDGNLYALKTADQYLYLNGNTAISAEODILVKG--TILKTSGLIA 772
Db 745 TNDVIRQYSYG-QPDDSYLALSDAENFLFNGNFLSTSKKEINVOGTRVIEYSGSN 803
Qy 773 TLERLQSFRLPEPLVQVLLTVGGEVFPKVKYTFVP----- 810
Db 804 AVERINSTNQEKEILLIQVLCV-GNLYNDVHYSFNIPLEKSDMTWTPYGMWECTYM 862
Qy 811 -----NDVDFSMQSKRE----- 822
Db 863 CQGLQRNITCTHKSDHSVVDKEDHLPLPSFVTQSCNTDCELRMLAKKSECSQCG 922
Qy 823 ----- 822
Db 923 QGYRTLDHCKMYSIHGGQVVDHYCDQDLKPLFQELCHGNCVTRNHYSEWQCSNS 982
Qy 823 -----RATTNIIQPLLA-----QWVLGWSSECSSTGAGW 853
Db 983 CGGGERESYCMNFGHRLADNECELSVTRBENGNFSCPSMAWSESECLVTCGKGT 1042
Qy 854 QRTVECR-DPSGQASATCNKALKPEAKCESQOLC 888
Db 1043 KORQVQCLNVDLSDGFNSSTKPBESLSPCELHTC 1078

```

```

ID AT15 MOUSE STANDARD; PRT; 562 AA.
AC P593B4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-15 (EC 3.4.24.-) (A disintegrin and metalloproteinase with
DE thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15) (Fragment).
GN ADAMTS15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=2238257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.T., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Bata S.S., Logucliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boeak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Paley J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rotherfield Y.S.N., Krzywinski M.I., Skalka U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: BC043308; AA043308.1; -
CC MGD: MGI:2449569; Adamts15.
DR InterPro: IPR006586; ADAM cysteine.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR001819; Pept_M10A_M12B.
DR InterPro: IPR006025; Pept_M2n_BS.
DR InterPro: IPR001590; Peptidase_M12B.
DR InterPro: IPR008084; TSP1.
DR InterPro: IPR008085; TSP_1.
DR Pfam: PF00090; TSP_1; 3.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SMO0608; ACR_1.
DR SMART: SMO0209; TSP1_3.
DR PROSITE: PS0215; ADAM_MPRO, PARTIAL.
DR PROSITE: PS00546; CYSTEINE_SWITCH, FALSE_NEG.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE: PS0214; DISINTEGRIN_2; FALSE_NEG.

```



Fri May 7 12:18:03 2004

us-09-989-687-4\_1.rsp

Page 20

Best Local Similarity 31.7%; Pred. No. 6.4e-67; Matches 329; Conservative 137; Mismatches 369; Indels 202; Gaps 42;

```
QY 3 PAPAAPMLPPLLLLLLLPLARG---APAPAAAGGQASLVETPLPGSAGELALHL 58
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 PPSRSP--APILRLPPLLLCALAPGAPGAPGRATEGRALDIVHPRVDAGSFLSYEL 62
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 -----SAF-----GKGFRLAPDDSLAPEKIE--RLGSGGRATGG 94
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 WPRALRKDVSVRDAPAFYELQYRGRELRLNLANOHLLAPGFVSETRRGGLGRA--- 119
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 95 ERGLR---GCFPGTNGEPE---SLAASLCRGLSGSFLDGEFTTQPGAGGSL-- 145
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 --HRAHTPACHLLGEVQ--DPELEGGLAALSACDGLKGVFLSNEDYFLEPLDSAPAPRG 176
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 146 -AOPHRLQRMGPAGARPLFRGFWEVETSGQROERGDHOEDSEESOEBAAGAEPP 204
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 HAQPRVYVK-----RQAPERLADGDSASASTGCVYPELSRRERWEGQRM 227
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 PLGATSRTKRF---VSEARFVETLLVADSMAPFYG-ADLQNHILLMSVAAATYKPS 259
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 F-----RLRLHOREVSEKKEVETLVADAKMVEYHGOQVSESYVLTIMMVAAGLFHDS 282
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 INKSNLWVVKLYVEDEKMPREVSDNGLLTLRFQWQRFQPSDRPHEHYDTAILLT 319
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 IGNPIHITIVRLVLEDEEDLKITTHADNTLKSFCMKQKSIKMGDAHPLHDTAILLT 342
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 320 RQNF--GQEGLCDTILGVADITICDPKSGSVIEDEGLQAHTLAHGLHLSMPHDS- 377
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 343 RKDLCAANRPECETLGLSHVAGMCPHRSINEDTLPLAFTVAHGLSHSGICHDGSG 402
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 378 KPCTRLFGPMGKH-HVNAPLFVHLNQLTPMSPCSAMLTLLDGGDCLLDAPGA-ALP 435
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 NDCE---FVGRPRFIMSPQLLYDAAPLWGRCSROYITRFLDRGMLCLDDEPACDIID 458
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 436 LPTGLPGMALYOLDQOCROIFPPDFRHCNPTSAQDVCAOLKCHTDGAEPLCHTKNGSLP 495
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 459 PPSVPPG--VLVDYSHOCRLQYASAFCE--MDNVCHTLWC---SVGTTCHSKLDA- 510
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 496 WADGTFCGPHLCSRGSCLEPEEVERPKPVVDGMAPMGPGSCRTGGGVQFSRECK 555
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 511 -VDGTRCGENKCKLSGECVPYG--FRPS-AVDGSGMSAMSIKSRSCGMVQSAERQCT 566
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 556 DEBPONGRYCLGRPAKQSCHTBECPPDGKSPRQCEKKNAYNTMDNGLLOWPKY 615
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 567 QPTPKYKGYCVGERKFRNLCLQAPGRPSFRHVQCSHEDAMLY---KGQLHTWVPV 623
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 616 AGVSPDRCKLECRARGSEFPVFEAKYIDGLCGPELTA--ICVRQCVKACDHYVDS 673
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 NDVNP---CELHCRANEYPAKLRDAVDDTFCYVARSRLCLINGICKNVGCDFEIDS 680
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 674 PKLKKCVGCGKNSCKRVSSGLFT-NYGRNDVITIPAGATNIDVKGRSHPGYQNDN 732
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 681 GAMEDRCGVCHNGSTCHTIVSGTEBAAGLYVDVGLIPAGAREIRIQE---VAAAN 735
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 733 YLAKTADGO-YLALNGNLAISAIEDILVKGTLIKYSSIATLERLOSFRPLPEPLTVOL 791
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 736 FLALNSEDPEKFLNGWTLIQ-WNGDYVAGTTFYAR--GNWENLTSPPGTKEPWIQY 793
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 792 -----LTVPG-EVFPKXYTFEVPN----- 811
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 794 PASRPGGSGRGVPRFSTLHGRSRPGSVSPGSVTERPGEPPAASISVSPSLKPNL 853
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 812 -----DYDFSMQSKERATTNIIQ 830
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 854 VAAVRRGQWQAPLGLGWRRLVLMGRLPTQLLFOESNGVAYEYTIHREAGGHDEVP 913
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 831 PLHAWVIGDMSESSCCGAG--WQRETVECRD-PSGQ-----ASATONKALKPEDA 880
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 914 PPFVS-WAYGPWTKCLVTCGRGKKGRHSPTCRGLVNSQGHNLQPLPAHCWATIGLEVCS 972
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 881 KP-----CEGQ---LCP 889
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Db 973 EPQFSICEMRLAIALCP 989

Search completed: May 7, 2004, 11:49:59  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:52:21 ; Search time 18 seconds  
(without alignments)  
2574.580 Million cell updates/sec

Title: US-09-989-687-4  
Perfect score: 890  
Sequence: 1 MFPAAPAPRWLPFLILLILL.....CNKALKPDAPKCESQCLPL 890

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size: 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database: SwissProt 42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	890	100.0	890	1	Q9U979 homo sapien
2	35	3.9	905	1	AT88_HUMAN
3	12	1.3	205	1	AT88_MOUSE
4	12	1.3	630	1	YK07_YEAST
5	12	1.3	837	1	AT84_RAT
6	12	1.3	860	1	AT86_HUMAN
7	12	1.3	967	1	AT81_RAT
8	12	1.3	968	1	AT81_MOUSE
9	11	1.2	232	1	FL3L_MOUSE
10	11	1.2	655	1	HGFA_HUMAN
11	10	1.1	152	1	NPW_PIG
12	10	1.1	165	1	NPW_HUMAN
13	10	1.1	185	1	NPW_MOUSE
14	10	1.1	245	1	AT84_BOVIN
15	10	1.1	245	1	CIQC_HUMAN
16	10	1.1	245	1	SHBG_RABIT
17	10	1.1	480	1	PRTP_HUMAN
18	10	1.1	492	1	MM1L_MOUSE
19	10	1.1	562	1	AT15_MOUSE
20	10	1.1	597	1	STIL_PANTR
21	10	1.1	741	1	PRO3_MOUSE
22	10	1.1	950	1	AT15_HUMAN
23	10	1.1	967	1	AT81_HUMAN
24	10	1.1	1007	1	CHC2_HUMAN
25	10	1.1	1013	1	PTPX_MACNE
26	10	1.1	1015	1	PTPX_HUMAN
27	10	1.1	1081	1	AT18_HUMAN
28	10	1.1	1277	1	CAML_FUGRU
29	10	1.1	1935	1	AT89_HUMAN
30	10	1.1	2923	1	CIR2_HUMAN
31	9	1.0	107	1	PRK2_RAT
32	9	1.0	114	1	NPW_MOUSE
33	9	1.0	128	1	PRK2_MOUSE

## ALIGNMENTS

RESULT 1	34	9	1.0	131	1	OREX_HUMAN	O43612 homo sapien
AT88_HUMAN	35	9	1.0	131	1	OREX_PIG	O77668 sus scrofa
ID	36	9	1.0	133	1	YJ23_YEAST	P47094 saccharomyc
AC	37	9	1.0	134	1	SECR_RAT	P11384 rattus norv
Q9UP79; Q9NZS0;	38	9	1.0	151	1	PAX2_MOUSE	Q9GKX3 mus musculu
DT	39	9	1.0	151	1	PAX2_RAT	Q9GKX3 rattus norv
16-OCT-2001 (Rel. 40, Created)	40	9	1.0	157	1	RNS2_PANTR	Q8GZP8 pan troglod
DT	41	9	1.0	161	1	UCN3_HUMAN	Q86Z83 homo sapien
15-MAR-2004 (Rel. 43, Last sequence update)	42	9	1.0	171	1	YCX7_YEAST	P25652 saccharomyc
DE	43	9	1.0	189	1	RMP2_MOUSE	Q9WUP0 mus musculu
ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2)	44	9	1.0	197	1	PTCA_MOUSE	Q64697 mus musculu
DE	45	9	1.0	196	1	TR22_MOUSE	Q9ER62 mus musculu
(METH-8).							
GN							
ADAMTS8 OR METH2.							
OS							
Homo sapiens (Human).							
OC							
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;							
CC							
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
ON							
NCBI TaxID=9606;							
RM							
SEQUENCE FROM N.A.							
RP							
TISSUE=Lung;							
RC							
MEDLINE=9367466; PubMed=10438512;							
RA							
Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,							
LA							
Lombardo M., Iruela-Arispe M.L., and METH-2 are members of a new							
RT							
family of proteins with angio-inhibitory activity."							
J. Biol. Chem. 274:23349-23357(1999).							
[2]							
SEQUENCE OF 195-440 FROM N.A.							
RP							
MEDLINE=20079168; PubMed=10610729;							
RX							
Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;							
RA							
"ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on							
RT							
mouse chromosome 9 and human chromosome 11."							
Genomics 62:312-315(1999).							
CC							
-1- FUNCTION: Has anti-angiogenic properties.							
CC							
-1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).							
CC							
-1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular							
CC							
matrix (By similarity).							
CC							
-1- TISSUE SPECIFICITY: Highly expressed in adult and fetal lung, lower							
CC							
expression in brain, placenta, heart, stomach and fetal brain and							
CC							
kidney.							
CC							
-1- DOMAIN: The spacer domain and the TSP type-1 domains are important							
CC							
for a tight interaction with the extracellular matrix.							
CC							
-1- PTM: The precursor is cleaved by a furin endopeptidase (By							
CC							
similarity).							
CC							
-1- SIMILARITY: Belongs to peptidase family M12B.							
CC							
-1- SIMILARITY: Contains 1 disintegrin-like domain.							
CC							
-1- SIMILARITY: Contains 2 TSP type-1 domains.							
CC							
This SWISS-PROT entry is copyright. It is produced through a collaboration							
CC							
between the Swiss Institute of Bioinformatics and the EMBL outstation -							
CC							
the European Bioinformatics Institute. There are no restrictions on its							
CC							
use by non-profit institutions as long as its content is in no way							
CC							
modified and this statement is not removed. Usage by and for commercial							
CC							
entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>							
CC							
or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).							
CC							
EMBL: AF060153; AF048081.1; -							
DR							
EMBL: AF175283; AF25806.1; -							



DR HSP: P34179; 11AG.  
 DR MEROPS; M12.226; -.  
 DR Genew; HGNC:224; ADAMTS8.  
 DR MIM; 605175; -.  
 DR GO; GO:0005179; F.integrin binding; TAS.  
 DR GO; GO:0008237; F.metalloproteinase activity; TAS.  
 DR GO; GO:0007345; P.embryogenesis and morphogenesis; TAS.  
 DR GO; GO:0008285; P.negative regulation of cell proliferation; TAS.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; tsp1; 2.  
 DR PRINTS; PRO1705; TSP1REPEAT.  
 DR SMART; SM00209; TSP1\_2.  
 DR PROSITE; PS00215; ADAM\_MERO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS00214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS00092; TSP1\_2.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR Hydrolase; Metalloproteinase; zinc; signal; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 27  
 FT PROPEP 28 214  
 FT CHAIN 215 890  
 FT DOMAIN 215 430  
 FT DOMAIN 439 526  
 FT DOMAIN 527 582  
 FT DOMAIN 584 690  
 FT DOMAIN 691 832  
 FT DOMAIN 834 889  
 FT DOMAIN 202 205  
 FT METAL 364 364  
 FT ACT\_SITE 365 365  
 FT METAL 368 368  
 FT METAL 374 374  
 FT CARBOHYD 345 345  
 FT CARBOHYD 401 401  
 FT CARBOHYD 466 466  
 FT CARBOHYD 491 491  
 FT CARBOHYD 600 600  
 FT CONFLICT 195 195  
 FT CONFLICT 413 440  
 FT SEQUENCE 890 AA; 96671 MM; 57D70EE03D5739D3 CRC64;  
 SQ  
 Query Match 100.0%; Score 890; DB 1; Length 890;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 FNQSPDRHPHYDTAILTRONFCQGEGLCDTLGVADIGITCDPNKSCSVTEDEGLAAH 360  
 DB 301 FNQSPDRHPHYDTAILTRONFCQGEGLCDTLGVADIGITCDPNKSCSVTEDEGLAAH 360  
 QY 361 TLAEHLGHTVSMPHDSKPCRTLPFGMGKHVMAFLFVHLNQTLPMSPCSMYLTLLDG 420  
 DB 361 TLAEHLGHTVSMPHDSKPCRTLPFGMGKHVMAFLFVHLNQTLPMSPCSMYLTLLDG 420  
 QY 421 GHGDCILDAFGAALPLPTGLPGRMALYOLDQCRTIPFPDRHCPNTSAQVCAQLMCHT 480  
 DB 421 GHGDCILDAFGAALPLPTGLPGRMALYOLDQCRTIPFPDRHCPNTSAQVCAQLMCHT 480  
 QY 481 DGAEPICHTNGSLPMDGTPCGPGLCSBGSCLPEEVEPRKPVVDGMAFMGMECS 540  
 DB 481 DGAEPICHTNGSLPMDGTPCGPGLCSBGSCLPEEVEPRKPVVDGMAFMGMECS 540  
 QY 541 RTGGGVQSHRECHDREPONGRYCLARRAYOSCHTEECPPPKSFRGQCKRYANN 600  
 DB 541 RTGGGVQSHRECHDREPONGRYCLARRAYOSCHTEECPPPKSFRGQCKRYANN 600  
 QY 601 YTDMDGNLLQWPKYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLGPEETLACVKG 660  
 DB 601 YTDMDGNLLQWPKYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLGPEETLACVKG 660  
 QY 661 QCVKAGCDVDSFRLDKGCGGKNSCKVSSGLPTNYGNDIVTIPAGNTINDYK 720  
 DB 661 QCVKAGCDVDSFRLDKGCGGKNSCKVSSGLPTNYGNDIVTIPAGNTINDYK 720  
 QY 721 QRSHPVQNDGNVLAALKTAGQYLLNGMLAISAEODILVKGITLKYSGLATLERLOSF 780  
 DB 721 QRSHPVQNDGNVLAALKTAGQYLLNGMLAISAEODILVKGITLKYSGLATLERLOSF 780  
 QY 781 RPLPELTQVLTLPGEVPPKVTTFPVNDVDSMOSKERTATNTIOQLLAQVNLG 840  
 DB 781 RPLPELTQVLTLPGEVPPKVTTFPVNDVDSMOSKERTATNTIOQLLAQVNLG 840  
 QY 841 DMSECSSTGAGMORRTVECRDPSCGASATCNKALKPEDAKCESQICPL 890  
 DB 841 DMSECSSTGAGMORRTVECRDPSCGASATCNKALKPEDAKCESQICPL 890

RESULT 2  
 ID AT88\_MOUSE STANDARD, PRT, 905 AA.  
 AC P57110;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAMTS-8 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).  
 GN ADAMTS8.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20079168; PubMed=10610729;  
 RA Georgiadis K.E., Hirokata S., Seldin M.F., Apte S.S.;  
 RT "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on mouse chromosome 9 and human chromosome 11.";  
 RL Genomics 62:312-315(1999).  
 CC -1- FUNCTION: Has anti-angiogenic properties (By similarity).  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed specifically in adult lung and heart and low expression during mouse development.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.  
 CC -1- PFM: The precursor is cleaved by a furin endopeptidase (By similarity).



CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF155282; AAF25805.1; -  
 CC HSSP; P34179; 11AG.  
 CC MEROPS; M12.226; -  
 CC MGD; MGI:1353468; Adams8.  
 CC InterPro; IPR001762; Disintegrin.  
 CC InterPro; IPR006025; Pept M Zn BS.  
 CC InterPro; IPR001590; Peptidase\_M12B.  
 CC InterPro; IPR000884; TSP1.  
 CC InterPro; IPR008085; TSP1.  
 CC Pfam; PF01421; Reprolysin; 1.  
 CC Pfam; PF00090; Tsp\_1; 2.  
 CC PRINTS; PR01705; TSP1REPEAT.  
 CC SMART; SM00209; TSP1; 2.  
 CC PROSITE; PS50215; ADAM\_MERO; 1.  
 CC PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 CC PROSITE; PS00214; DISINTEGRIN\_2; FALSE\_NEG.  
 CC PROSITE; PS00392; TSP1; 2.  
 CC PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 CC Repeat; Extracellular matrix; Heparin-binding.  
 CC SIGNAL  
 CC PROPEP 1 228  
 CC CHAIN 229 905  
 CC DOMAIN 229 452  
 CC DOMAIN 453 541  
 CC DOMAIN 542 597  
 CC DOMAIN 599 705  
 CC DOMAIN 706 847  
 CC DOMAIN 848 904  
 CC METL 378 378  
 CC ACT SITE 379 379  
 CC METL 382 382  
 CC METL 388 388  
 CC CARBOHYD 415 415  
 CC CARBOHYD 480 480  
 CC CARBOHYD 506 506  
 CC CARBOHYD 615 615  
 CC SEQUENCE 905 AA; 98879 MW; 124D4132B33A0CAE CRC64;  
 Query Match 3.9%; Score 35; DB 1; Length 905;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-22;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 617 GVSPDRCKLFCRARGSEFFVFAKYIDGTLCP 651  
 DB 632 GVSPDRCKLFCRARGSEFFVFAKYIDGTLCP 666  
 ID YK07\_YEAST STANDARD; PRT; 205 AA.  
 AC P36061;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Hypothetical 22.7 kDa protein in SDH-C1M5/YT3 intergenic region.  
 GN YKL147C OR YKL601.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NCBI\_TaxID=4932;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=8288c;  
 RX MEDLINE=84378720; PubMed=8091859;  
 RA Vanderbol M., Bolle P.-A., Dion C., Portetelle D., Hlager F.  
 RT LAP loci of chromosome XI of *Saccharomyces cerevisiae*.  
 RL Yeast 10:535-540(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z26877; CA81507.1; -  
 CC EMBL; Z28146; CA81987.1; -  
 CC PIR; S37804; S37804.  
 CC GenOnline; 139903; -  
 CC SGD; S0001630; YKL147C.  
 CC Hypothetical protein.  
 CC SEQUENCE 205 AA; 22673 MW; 01F162B361D9DCE CRC64;  
 Query Match 1.3%; Score 12; DB 1; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 0.0078;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 13 FLLLLLLLLPL 24  
 DB 73 FLLLLLLLLPL 84  
 ID AT54\_RAT STANDARD; PRT; 630 AA.  
 AC Q9ESP7; Q9ESP6; Q9ESP8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-4 precursor (EC 3.4.24.82) (A disintegrin and metalloproteinase  
 DE with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)  
 DE (Fragment).  
 GN ADAMTS4.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mistar; Tissue=Brain;  
 RX MEDLINE=20415831; PubMed=10961658;  
 RA Satch K., Suzuki N., Yokota H.,  
 RT "ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin  
 RT motifs) is transcriptionally induced in beta-amyloid treated rat  
 RT astrocytes."  
 RL Neurosci. Lett. 289:177-180(2000).  
 CC -----  
 CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover. May play an important role in the  
 CC destruction of aggrecan in arthritic diseases. Cleaves aggrecan at  
 CC the 392-Glu-Ala-393 site.  
 CC -1- CATALYTIC ACTIVITY: GlutamyI endopeptidase; bonds cleaved include  
 CC 370-Thr-Glu-Gly-Glu-Ala-His-Gly-Ser-377 in the interglobular  
 CC domain of mammalian aggrecan.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Only expressed in brain.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase.  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.

CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB042272; BAB16474.1; -  
 CC EMBL; AB042273; BAB16475.1; -  
 CC EMBL; AB042273; BAB16475.1; -  
 CC InterPro: IPR001762; Disintegrin.  
 CC InterPro: IPR006025; Pept M Zn BS.  
 CC InterPro: IPR001590; Peptidase\_M1B.  
 CC InterPro: IPR000884; TSP1.  
 CC InterPro: IPR008085; TSP1.  
 CC Pfam: PF01421; Reprolysin; 1.  
 CC Pfam: PF00050; TSP1; 1.  
 CC PRINTS; PR01705; TSP1REPEAT.  
 CC SMART; SM00209; TSP1; 1.  
 CC PROSITE; PS02025; ADAM\_NEPRO; 1.  
 CC PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 CC PROSITE; PS00214; DISINTEGRIN\_2; FALSE\_NEG.  
 CC PROSITE; PS00092; TSP1; 1.  
 CC PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 CC HydroLase; Metalloprotease; Zinc; Glycoprotein; Zymogen;  
 CC Extracellular matrix.  
 CC FT NON TER 1 1  
 CC FT PROPEP <1 5  
 CC FT CHAIN 6 630 BY SIMILARITY.  
 CC FT DOMAIN 5 232 ADAMTS-4.  
 CC FT DOMAIN 6 303 METALLOPROTEASE.  
 CC FT DOMAIN 223 303 DISINTEGRIN-LIKE.  
 CC FT DOMAIN 313 368 TSP TYPE-1.  
 CC FT DOMAIN 368 478 CYS-RICH.  
 CC FT DOMAIN 479 630 SPACER.  
 CC FT DOMAIN 40 45 POLY-ALA.  
 CC FT METAL 154 154 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC FT ACT SITE 155 155 BY SIMILARITY.  
 CC FT METAL 158 158 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC FT METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SQ SEQUENCE 630 AA; 68384 MW; 634428753167C7EF CRC64;  
 CC  
 CC Query Match 1.3%; Score 12; DB 1; Length 630;  
 CC Best Local Similarity 100.0%; Pred. No. 0.019;  
 CC Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 539 GSRTGGGVQFS 550  
 CC DB 325 GSRTGGGVQFS 336  
 CC  
 CC RESULT 5  
 CC ATSA4 HUMAN  
 CC ID 075173; OSUN83; STANDARD; PRT; 837 AA.  
 CC AC 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 CC DE ADAMTS-4 precursor (EC 3.4.24.82) (A disintegrin and metalloproteinase  
 CC with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)  
 CC (ADMP-1).  
 CC GN ADAMTS4 OR KIAA0688.  
 CC OS Homo sapiens (human).  
 CC CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CC OX NCBI\_Taxid=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.

CC TISSUE=Brain;  
 CC MEDLINE=98403880; PubMed=9734811;  
 CC Iahkawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,  
 CC Kohani H., Nomura N., Ohara O.;  
 CC "Prediction of the coding sequences of unidentified human genes. X.  
 CC The complete sequences of 100 new cDNA clones from brain which can  
 CC code for large proteins in vitro.";  
 CC DNA Res. 5:169-176 (1998).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE=99286303; PubMed=10356395;  
 CC Torrorella M.D., Burn T.C., Pratta M.A., Abbazade I., Hollis J.M.,  
 CC Liu R.-Q., Rosenthal S.A., Copeland R.A., Decicco C.F., Wyman R.,  
 CC Rockwell A., Yang F., Duke J.B., Solomon K., George H., Bruckner R.,  
 CC Nagase H., Itoh Y., Ellis D.M., Rose H., Mawell B.H., Murphy K.,  
 CC Hillman M.C., Jr., Hollis G.F., Newton R.C., Magolda R.L.,  
 CC Tzaskos J.M., Arner E.C.;  
 CC "Purification and cloning of aggrecanase-1: a member of the ADAMTS  
 CC family of proteases.";  
 CC Science 284:1664-1666 (1999).  
 CC [3]  
 CC SEQUENCE FROM N.A.  
 CC Sawaji Y., Nagase H., Sakatvala J., Clark A.R.;  
 CC "ADAMTS-4 genomic locus.";  
 CC submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
 CC [4]  
 CC PARTIAL SEQUENCE AND CHARACTERIZATION.  
 CC MEDLINE=20400518; PubMed=10827174;  
 CC Torrorella M.D., Pratta M.A., Liu R.-Q., Abbazade I., Rose H.,  
 CC Burn T.C., Arner E.C.;  
 CC "The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for  
 CC aggrecan substrate recognition and cleavage.";  
 CC J. Biol. Chem. 275:25791-25797 (2000).  
 CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover. May play an important role in the  
 CC destruction of aggrecan in arthritic diseases. Could also be a  
 CC critical factor in the exacerbation of neurodegeneration in 393  
 CC Alzheimer's disease. Cleaves aggrecan at the 392-Glu-Ala-393  
 CC site.  
 CC -1- CATALYTIC ACTIVITY: glutamyl endopeptidase; bonds cleaved include  
 CC 370-Thr-Glu-Gly-Glu-Ala-Arg-Gly-Ser-377 in the interglobular  
 CC domain of mammalian aggrecan.  
 CC -1- COFACTOR: Birds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed in brain, lung and heart. Expressed  
 CC at very low level in placenta and skeletal muscles.  
 CC -1- INDUCTION: By interleukin-1.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase.  
 CC -1- SIMILARITY: Belongs to peptidase family M1B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC -1- CAUTION: Has sometimes been referred to as ADAMTS2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB014588; BAA31663.1; -  
 CC EMBL; AF148213; AAD41494.1; -  
 CC EMBL; AY044847; AAL02262.1; -  
 CC PIR; T00355; T00355.  
 CC DR MEROPS; M2.221; -  
 CC DR Gene; HGNC:220; ADAMTS4.  
 CC MIM; 603876; -  
 CC DR GO; GO:0008237; F:metallopeptidase activity; TAs  
 CC GO; GO:0006508; F:proteolysis and peptidolysis; TAs.

DR GO:0001501; P.skeletal development; TAS.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M2n\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01421; Repolysin; 1.  
 DR Pfam; PF00090; tsp.1; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM0209; TSP1; 1.  
 DR PROSITE; PS50215; ADAM\_MERPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Extracellular matrix.  
 FT SIGNAL 1 51 POTENTIAL.  
 FT PROPEP 52 212  
 FT CHAIN 213 837 ADAMTS-4.  
 FT DOMAIN 213 436 METALLOPROTEASE.  
 FT DOMAIN 437 519 DISINTEGRIN-LIKE.  
 FT DOMAIN 520 575 TSP TYPE-1.  
 FT DOMAIN 577 685 CYS-RICH.  
 FT DOMAIN 686 837 SPACER.  
 FT DOMAIN 247 252 POLY-ALA.  
 FT SITE 194 194 CYSTEINE SWITCH (POTENTIAL).  
 FT METAL 361 361 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 362 362 BY SIMILARITY.  
 FT METAL 365 365 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 371 371 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 77 77 A -> T (IN REF. 1).  
 FT CONFLICT 626 626 R -> Q (IN REF. 3).  
 FT CONFLICT 682 682 G -> R (IN REF. 3).  
 SQ SEQUENCE 837 AA; 90224 MW; 5DF9C9AC137DF41F CRC64;

Query Match 1.3%; Score 12; DB 1; Length 837;  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGAGVGS 550  
 DB 532 CSRTCGGAGVGS 543

RESULT 6  
 A1S6\_HUMAN STANDARD; PRT; 860 AA.  
 ID A1S6\_HUMAN  
 AC Q9URF5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAMTS-6 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 6) (ADAM-TS 6) (ADAM-TS6).  
 GN ADAMTS6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99395124; PubMed=10464288;  
 RA Hukariainen T.U., Hirohata S., Seidlin M.F., Agre S.S.;  
 RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of  
 RT zinc metalloproteases.";  
 RL J. Biol. Chem. 274:25555-25563(1999).  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed at low levels in placenta and barely  
 CC detectable in a number of other tissues.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important

CC for a tight interaction with the extracellular matrix.  
 CC -1- PFM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF140674; MAB56357.1; -.  
 CC HSSP; P34179; I1AG.  
 CC MEROPS; M12.230; -.  
 CC Genew; HGNC:222; ADAMTS6.  
 CC MIM; 605008; -.  
 DR GO:0008237; F:metallopeptidase activity; TAS.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M2n\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Pep\_M12B\_Proped; 1.  
 DR Pfam; PF00491; Repolysin; 1.  
 DR Pfam; PF00090; tsp.1; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM0209; TSP1; 1.  
 DR PROSITE; PS50215; ADAM\_MERPRO; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 244 BY SIMILARITY.  
 FT CHAIN 245 860 ADAMTS-6.  
 FT DOMAIN 245 452 METALLOPROTEASE.  
 FT DOMAIN 453 509 DISINTEGRIN-LIKE.  
 FT DOMAIN 510 565 TSP TYPE-1.1.  
 FT DOMAIN 567 668 CYS-RICH.  
 FT DOMAIN 669 795 SPACER.  
 FT DOMAIN 796 852 TSP TYPE-1.2.  
 FT DOMAIN 852 852 POLY-ARG.  
 FT DOMAIN 68 71 POLY-GLY.  
 FT METAL 403 403 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 404 404 BY SIMILARITY.  
 FT METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHYD 99 99 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 172 172 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 222 222 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 676 676 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 843 843 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 860 AA; 97098 MW; E57213015DEC82C5 CRC64;

Query Match 1.3%; Score 12; DB 1; Length 860;  
 Best Local Similarity 100.0%; Pred. No. 0.025;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WBCSRTCGGAV 547  
 DB 519 WBCSRTCGGAV 530

RESULT 7  
 A1S1\_RAT

ID ATSL RAT STANDARD; PRT; 967 AA.  
 AC 09M001; 09M011; Rel. 40, Created  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).  
 GN ADAMTS1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NC NCB1\_taxonomy=10116;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RA Liu X., Tu Y., Yan T., Johnstone E.M., Stephenson D.T., Clemens J.A.,  
 RA Little S.P.;  
 RT "Induction of a disintegrin and metalloproteinase with the  
 thrombospondin type I motif (ADAMTS).";  
 RT Submitted (May-1999) to the EMBL/Genbank/DBJ databases.  
 RN (2)  
 RP SEQUENCE OF 18-967 FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
 RX MEDLINE=20304099; PubMed=10847486;  
 RA Diamantis I., Luechi M., Hoessli M., Reichen J.;  
 RT "Cloning of the rat ADAMTS-1 gene and its down regulation in  
 endothelial cells in cirrhotic rats.";  
 RT Liver 20:165-172(2000).  
 CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 involved in its turnover. Has angiogenic inhibitor activity (By  
 similarity). Active metalloproteinase, which may be associated with  
 various inflammatory processes as well as development of cancer  
 cachexia. May play a critical role in follicular rupture (By  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-|-Leu-1684  
 site, within the chondroitin sulfate attachment domain.  
 CC -1- CORRECTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 matrix (By similarity).  
 CC -1- INDUCTION: Down-regulated in endothelial cells derived from  
 cirrhotic liver.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 for a tight interaction with the extracellular matrix.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By  
 similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 3 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF149118; AAC34012.1; -  
 DR EMBL; AF304446; AAG29823.1; -  
 DR MEROPS; M12.222; -  
 DR InterPro; IPR006586; ADAM\_cysteine.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Rep\_M12B\_Dipep; 1.  
 DR Pfam; PF01421; RepProlysin; 1.  
 DR Pfam; PF00090; TSP1; 3.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 3.

DR PROSITE; PS50215; ADAM\_MEROPS; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix; Heparin-binding.  
 FT SIGNAL; 1 54  
 FT PROPEP; 55 252  
 FT CHAIN; 253 475  
 FT DOMAIN; 253 475  
 FT DOMAIN; 476 558  
 FT DOMAIN; 559 614  
 FT DOMAIN; 616 724  
 FT DOMAIN; 725 857  
 FT DOMAIN; 854 910  
 FT DOMAIN; 911 967  
 FT DOMAIN; 194 198  
 FT SITE; 205 205  
 FT METAL; 401 401  
 FT ACT\_SITE; 402 402  
 FT METAL; 405 405  
 FT METAL; 411 411  
 FT CARBOHYD; 547 547  
 FT CARBOHYD; 720 720  
 FT CARBOHYD; 764 764  
 FT CARBOHYD; 782 782  
 FT CARBOHYD; 945 945  
 FT CONFLICT; 21 21  
 FT CONFLICT; 26 31  
 FT CONFLICT; 46 49  
 FT CONFLICT; 72 72  
 FT CONFLICT; 79 79  
 FT CONFLICT; 249 249  
 FT CONFLICT; 262 265  
 FT CONFLICT; 607 607  
 FT CONFLICT; 936 936  
 FT CONFLICT; 962 962  
 SQ SEQUENCE 967 AA; 105705 MW; P9JC864F6CDB4CF CRC64;  
 Query Match 1.34; Score 12; DB 1; Length 967;  
 Best Local Similarity 100.0%; Pred. No. 0.027;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 706 DIVITPAGATNI 717  
 DB 740 DIVITPAGATNI 751  
 RESULT 8  
 ATSL MOUSE STANDARD; PRT; 968 AA.  
 ID ATSL MOUSE  
 AC P97857; 054768;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).  
 GN ADAMTS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCB1\_taxonomy=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ;  
 RX MEDLINE=98110583; PubMed=9441751;  
 RA Kuno K., Iizasa H., Ohno S., Matsushima K.;  
 RT "The exon/intron organization and chromosome mapping of the mouse  
 ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";  
 RT Genomics 46:466-471(1997).  
 RN (2)  
 RP SEQUENCE FROM N.A.

RX MEDLINE=97150761; PubMed=8995297;  
 RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,  
 RA Matsushima K.;  
 RT "Molecular cloning of a gene encoding a new type of metalloproteinase-  
 RT disintegrin family protein with thrombospondin motifs as an  
 RT inflammation associated gene.";  
 RL J. Biol. Chem. 272:556-562(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Limb, and Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Struhsberg R.U., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carrinci P., Prange C.,  
 RA Rana S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Plakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP CHARACTERIZATION AND MUTAGENESIS OF GLU-403.  
 RX MEDLINE=99303657; PubMed=10373507;  
 RA Kuno K., Terahashima Y., Matsushima K.;  
 RT "ADAMTS-1 is an active metalloproteinase associated with the  
 RT extracellular matrix.";  
 RL J. Biol. Chem. 274:18821-18826(1999).  
 RN [5]  
 RP FUNCTION.  
 RX MEDLINE=20389568; PubMed=10930576;  
 RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,  
 RA Ohno H., Matsushima K.;  
 RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";  
 RL FEBS Lett. 478:241-245(2000).  
 RN [6]  
 RP FUNCTION, AND INDUCTION.  
 RX MEDLINE=20243757; PubMed=10781075;  
 RA Robker R.L., Russell D.V., Espey L.L., Lydon J.P., O'Malley B.W.,  
 RA Richards U.S.;  
 RT "Progestosterone-regulated genes in the ovulation process: ADAMTS-1 and  
 RT cathepsin L proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).  
 CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover. Has antigenic inhibitor activity (BY  
 CC similarity). Active metalloproteinase, which may be associated with  
 CC various inflammatory processes as well as development of cancer  
 CC cachexia. May play a critical role in follicular rupture (BY  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-1720-1692  
 CC site, within the chondroitin sulfate attachment domain.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (BY similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix.  
 CC -1- INDUCTION: Induced in vitro in colon adenocarcinoma cells by  
 CC interleukin-1, or in vivo in kidney and heart by  
 CC lipopolysaccharide. Also induced by LH stimulation in granulosa  
 CC cells of preovulatory follicles.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domain are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase.  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.

CC -1- SIMILARITY: Contains 3 TSP type-1 domain.  
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to a  
 CC frameshift in position 7.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC DR EMBL; AB001735; BA24501.1; ALT INT.  
 CC DR EMBL; D67076; BA11088.1; ALT FRAM.  
 CC DR EMBL; BC040382; AA40382.1; -  
 CC DR EMBL; BC050834; AA50834.1; -  
 CC DR MEROPS; M12.222; -  
 CC DR MGD; MGI:109249; Adamts1.  
 CC DR InterPro; IPR006586; ADAM\_cysteine.  
 CC DR InterPro; IPR001762; Disintegrin.  
 CC DR InterPro; IPR006025; Pept\_M\_zn\_BS.  
 CC DR InterPro; IPR001590; Peptidase\_M12B.  
 CC DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 CC DR InterPro; IPR000884; TSP1.  
 CC DR InterPro; IPR008085; TSP1.  
 CC DR Pfam; PF01562; TSP1.  
 CC DR Pfam; PF01421; Reprolysin.  
 CC DR Pfam; PF00090; TSP1.  
 CC DR PRINTS; PR01705; TSP1REPEAT.  
 CC DR SMART; SMO0209; TSP1.  
 CC DR SMART; SMO0608; ACR1.  
 CC DR PROSITE; PS0215; ADAM\_MERO. 1.  
 CC DR PROSITE; PS00427; DISINTEGRIN\_1. FALSE NEG.  
 CC DR PROSITE; PS0214; DISINTEGRIN\_2. FALSE NEG.  
 CC DR PROSITE; PS0092; TSP1. 3.  
 CC DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 CC DR Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 CC KW Repeat; Extracellular matrix; Heparin-binding.  
 CC FT SIGNAL; 1  
 CC FT PROPEP; 48  
 CC FT CHAIN; 254  
 CC FT DOMAIN; 254  
 CC FT DOMAIN; 477  
 CC FT DOMAIN; 560  
 CC FT DOMAIN; 618  
 CC FT DOMAIN; 725  
 CC FT DOMAIN; 726  
 CC FT DOMAIN; 855  
 CC FT DOMAIN; 912  
 CC FT DOMAIN; 919  
 CC FT SIBR; 206  
 CC FT METBL; 402  
 CC FT ACT SITE; 403  
 CC FT METBL; 406  
 CC FT METBL; 412  
 CC FT CARBOHYD; 548  
 CC FT CARBOHYD; 721  
 CC FT CARBOHYD; 765  
 CC FT CARBOHYD; 783  
 CC FT CARBOHYD; 946  
 CC FT MUTAGEN; 403  
 CC FT CONFLICT; 335  
 CC FT CONFLICT; 425  
 CC FT SEQUENCE; 968 AA; 105841 MW; 42EBDA549F86C1 CRC64;  
 CC -----  
 CC Query Match 1.3%; Score 12; DB 1; Length 968;  
 CC Best Local Similarity 100.0%; Pred. No. 0.027; 0; Indels 0; Gaps 0;  
 CC Matches 12; Conservative 0; Mismatches 0;  
 CC -----  
 CC QY 706 DIVTIPAGATNI 717  
 CC |||||  
 CC DB 741 DIVTIPAGATNI 752

```

RESULT 9
F3L3_MOUSE STANDARD; PRT; 232 AA.
ID F3L3_MOUSE
AC P49772; 064085;
AD 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SL cytokine precursor [fms-related tyrosine kinase 3 ligand] (flt3
ligand) (flt3l).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_faxId=100930;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94195428; PubMed=8145851;
RA Hannun C., Culpepper J., McClanahan T., Zurawski S.,
Bada G., Katselstein R., Hudak S., Wagner J., Matson J., Luh J.,
Duran J.F., Maratea R., Peterson D., Menon S., Shanfield A.,
Zlotnik A., Rosnet O., Nemikawa R., Rennick D., Roncarolo M.G.,
"ligand for FLT3/FLK2 receptor tyrosine kinase regulates growth of
hematopoietic stem cells and is encoded by variant RNAs.";
RL Nature 368:643-648(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SW/J;
RX MEDLINE=94084791; PubMed=7505204;
RA Lyman S.D., James L., Vanden Bos T., Devries P., Brasel K.,
Gliniak B., Hollingsworth L.T., Fitch K.S., McKenna H.C.,
Splet R.R., Fletcher F.A., Marksovsky E., Farrar T.,
Foxworth D., Williams D.E., Beckmann M.P.;
"Cloning of a proliferative factor for primitive hematopoietic cells."
Cell 75:1157-1167(1993).
RN [3]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=6032581; PubMed=7568777;
RA Lyman S.D., Stocking K., Davison B., Fletcher P., Johnson L.,
Beckor S.;
"Structural analysis of human and murine flt3 ligand genomic loci.";
RL Oncogene 11:1165-1172(1995).
RN [4]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=95124710; PubMed=7824267;
RA Lyman S.D., James L., Beckor S., Downey H., de Vries P.,
Brasel K., Stocking K., Beckmann M.P., Copeland N.G.,
Cleveland L.S.;
"Identification of soluble and membrane-bound isoforms of the murine
flt3 ligand generated by alternative splicing of mRNAs.";
RL Oncogene 10:149-157(1995).
RN [5]
RP SEQUENCE FROM N.A.
RA McClanahan T., Culpepper J., Campbell D., Wagner J.,
Franz-Bacon K., Matson J., Tsai S., Luh J., Guinardes M.J.,
Wattel M.-G., Rosnet O., Birnbaum D., Hannum C.;
Submitted (APR-1996) to the EMBL/Genbank/DBD databases.
RL -1- FUNCTION: Stimulates the proliferation of early hematopoietic
cells. Synergizes well with a number of other colony stimulating
factors and interleukins.
CC -1- SUBUNIT: Homodimer (soluble isoform) (By similarity).
CC -1- STRUCTURAL LOCATOR: Type I membrane protein. Two soluble
isoforms are also produced by alternative splicing. One of which
is biologically active, while the other, isoform
4/Bdelta1a, is inactive.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-alternative product(s); Named isoforms=4;
CC Name=1; Synonyms=6C;
CC IsoId=P49772-1; Sequence=Displayed;
CC Name=2; Synonyms=5H;
CC IsoId=P49772-2; Sequence=VSP_004253;
CC Name=3; Synonyms=56;

```

[illegible]

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RX NCBI\_TaxID=9606;  
 RP [1]  
 RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Liver, and Serum; PubMed=768365;  
 RX MEDLINE=93252878; PubMed=768365;  
 RA Miyazawa K., Shimamura T., Kitamura A., Kondo J., Morimoto Y.,  
 RA Kitamura N.;  
 RT "Molecular cloning and sequence analysis of the cDNA for a human  
 RT serine protease responsible for activation of hepatocyte growth  
 RT factor. Structural similarity of the protease precursor to blood  
 RT coagulation factor XII.";  
 RL J. Biol. Chem. 268:10024-10028 (1993).  
 RP [2]  
 RP SEQUENCE OF 40-655 FROM N.A.  
 RA Zhao S., Odell C.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: Activates hepatocyte growth factor (HGF) by  
 CC converting it from a single chain to a heterodimeric form.  
 CC -1 SUBUNIT: Dimer of a short chain and a long chain linked by a  
 CC disulfide bond.  
 CC -1 SUBCELLULAR LOCATION: Secreted as an inactive single-chain  
 CC precursor and is then activated to a heterodimeric form.  
 CC -1 TISSUE SPECIFICITY: Liver.  
 CC -1 SIMILARITY: Belongs to peptidase family S1.  
 CC -1 SIMILARITY: Contains 2 EGF-like domains.  
 CC -1 SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -1 SIMILARITY: Contains 1 Kringle domain.  
 CC -1 CAUTION: It is uncertain whether Met-1 is the initiator.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: D14012; BA03113.1;  
 DR EMBL: Z69923; -; NOT\_ANNOTATED\_CDS.  
 DR PTR: A46688; A46688.  
 DR HSRP: P00763; IDPO.  
 DR MEROPS: S01.228;  
 DR Genew: HENC:4894; HGFPAC.  
 DR MIM: 604552;  
 DR GO: GO:0005576; C:extracellular; TAS.  
 DR GO: GO:0004252; F:serine-type endopeptidase activity; TAS.  
 DR GO: GO:0006508; P:proteolysis and peptidolysis; TAS.  
 DR InterPro: IPR009042; Cys\_Ser\_trypsin.  
 DR InterPro: IPR006209; EGF\_like.  
 DR InterPro: IPR000083; Fibrinctn.  
 DR InterPro: IPR000562; FN\_Type\_II.  
 DR InterPro: IPR006210; IEGF.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR001254; Peptidase\_S1.  
 DR InterPro: IPR001314; Peptidase\_S1A.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00039; fn1; 1.  
 DR Pfam: PF00040; fn2; 1.  
 DR Pfam: PF00051; Kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00013; ENTPEP1.  
 DR PRINTS: PR00018; KRINGLE.  
 DR ProDom: PD000995; FN\_Type\_II; 1.  
 DR ProDom: PD000395; Kringle; 1.  
 DR SMART: SM00181; EGF; 2.  
 DR SMART: SM00056; FN1; 1.  
 DR SMART: SM00059; FN2; 1.  
 DR SMART: SM00130; KR; 1.  
 DR SMART: SM00020; Tryp\_Spe; 1.

DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 2.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;  
 KW EGF-like domain; Repeat; Zymogen.  
 FT SIGNAL 1 30  
 FT PROPEP 31 372  
 FT CHAIN 373 407  
 FT CHAIN 408 655  
 FT DOMAIN 108 148  
 FT DOMAIN 160 198  
 FT DOMAIN 200 240  
 FT DOMAIN 241 279  
 FT DOMAIN 286 367  
 FT DOMAIN 408 655  
 FT ACT\_SITE 447 447  
 FT ACT\_SITE 497 497  
 FT ACT\_SITE 598 598  
 FT DISULFID 108 133  
 FT DISULFID 122 148  
 FT DISULFID 164 175  
 FT DISULFID 169 186  
 FT DISULFID 188 197  
 FT DISULFID 202 230  
 FT DISULFID 228 237  
 FT DISULFID 245 256  
 FT DISULFID 250 267  
 FT DISULFID 278 278  
 FT DISULFID 286 367  
 FT DISULFID 307 349  
 FT DISULFID 338 362  
 FT DISULFID 394 521  
 FT DISULFID 432 448  
 FT DISULFID 440 510  
 FT DISULFID 535 604  
 FT DISULFID 567 583  
 FT DISULFID 594 622  
 FT CARBOHYD 48 48  
 FT CARBOHYD 290 290  
 FT CARBOHYD 468 468  
 FT CARBOHYD 492 492  
 FT CARBOHYD 546 546  
 FT CONFLICT 644 644  
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B86ED7 CR664;  
 Query Match 1.2%; Score 11; DB 1; Length 655;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Cy 12 PLLLLLLLLL 22  
 Db 18 PLLLLLLLLL 28  
 RESULT 11  
 NPW\_PIG STANDARD; PRT; 152 AA.  
 AC OSW135;  
 ID 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Neuropeptide W precursor (Preproprotein W8) (PPI8) [Contains:  
 DE Neuropeptide W-23 (NPW23) (L8); Neuropeptide W-30 (NPW30) (L8C)].  
 GN NPW OR PPI8.



OS Sus scrofa (pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OC NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 33-52, AND SYNTHESIS OF NPW23 AND  
 RP NPW30.  
 RX MEDLINE=22229377; PubMed=12130646;  
 RA Shimomura Y., Harada M., Goto M., Sugo T., Matsumoto Y., Abe M.,  
 RA Watanabe T., Asami T., Kitada C., Mori M., Onda H., Fujino M.;  
 RT "Identification of neuropeptide W as the endogenous ligand for orphan  
 RT G-protein-coupled receptor GPR7 and GPR8.";  
 RL J. Biol. Chem. 277:35826-35832(2002).  
 CC -1- FUNCTION: Plays a regulatory role in the organization of  
 CC neuroendocrine signals accessing the anterior pituitary gland.  
 CC stimulates water drinking and food intake. May play a role in the  
 CC hypothalamic response to stress (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the neuropeptide B/W family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL, AB084277; BAC07173.1; -  
 KW Neuropeptide; Cleavage on pair of basic residues; Signal.  
 FT SIGNAL 1 32  
 FT PEPTIDE 33 55 NEUROPEPTIDE W-23.  
 FT PROPEP 65 152 NEUROPEPTIDE W-30.  
 FT PROPEP 152 16347 MW; 454284DD44806784 CRC64;  
 SQ SEQUENCE 152 AA; 16347 MW; 454284DD44806784 CRC64;  
 Query Match 1.1%; Score 10; DB 1; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.34; Mismatches 0; Indels 0;  
 Matches 10; Conservative 0; Mismatches 0;  
 QY 15 LLLLLLLPL 24  
 Db 19 LLLLLLLPL 28  
 RESULT 12  
 NPW\_HUMAN STANDARD; PRT; 165 AA.  
 AC Q8N729;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Neuropeptide W precursor (Preproprotein 18) (pp18) [Contains:  
 DE Neuropeptide W-23 (NPW23) (HL8); Neuropeptide W-30 (NPW30) (HL8C)].  
 GN NPW OR PP18 OR PP18.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND SYNTHESIS OF NPW23 AND  
 RP NPW30.  
 RX MEDLINE=22229377; PubMed=12130646;  
 RA Shimomura Y., Harada M., Goto M., Sugo T., Matsumoto Y., Abe M.,  
 RA Watanabe T., Asami T., Kitada C., Mori M., Onda H., Fujino M.;  
 RT "Identification of neuropeptide W as the endogenous ligand for orphan  
 RT G-protein-coupled receptors GPR7 and GPR8.";  
 RL J. Biol. Chem. 277:35826-35832(2002).  
 CC -1- FUNCTION: Plays a regulatory role in the organization of  
 CC neuroendocrine signals accessing the anterior pituitary gland.  
 CC stimulates water drinking and food intake. May play a role in the  
 CC hypothalamic response to stress (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the neuropeptide B/W family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL, AB084276; BAC07172.1; ALT\_INIT.  
 KW Neuropeptide; Cleavage on pair of basic residues; Signal.  
 FT SIGNAL 1 32  
 FT PEPTIDE 33 55 NEUROPEPTIDE W-23.  
 FT PROPEP 65 165 NEUROPEPTIDE W-30.  
 FT PROPEP 165 18048 MW; F426AF457F6B604 CRC64;  
 SQ SEQUENCE 165 AA; 18048 MW; F426AF457F6B604 CRC64;  
 Query Match 1.1%; Score 10; DB 1; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.36;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0;  
 QY 15 LLLLLLLPL 24  
 Db 19 LLLLLLLPL 28

RT "Identification of natural ligands for the orphan G protein-coupled  
 RT receptors GPR7 and GPR8.";  
 RL J. Biol. Chem. 278:176-783(2003).  
 CC -1- FUNCTION: Plays a regulatory role in the organization of  
 CC neuroendocrine signals accessing the anterior pituitary gland.  
 CC stimulates water drinking and food intake. May play a role in the  
 CC hypothalamic response to stress (By similarity). NPW23 activates  
 CC GPR7 and GPR8 more efficiently than NPW30.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Detected at high levels in the substantia  
 CC nigra, fetal kidney and trachea; at lower levels in testis,  
 CC uterus, ovary and placenta. Not detectable in many regions of the  
 CC central nervous system. Also detected at high levels in  
 CC lymphoblastic leukemia and colorectal adenocarcinoma.  
 CC -1- SIMILARITY: Belongs to the neuropeptide B/W family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL, AB084276; BAC07172.1; ALT\_INIT.  
 KW Neuropeptide; Cleavage on pair of basic residues; Signal.  
 FT SIGNAL 1 32  
 FT PEPTIDE 33 55 NEUROPEPTIDE W-23.  
 FT PROPEP 65 165 NEUROPEPTIDE W-30.  
 FT PROPEP 165 18048 MW; F426AF457F6B604 CRC64;  
 SQ SEQUENCE 165 AA; 18048 MW; F426AF457F6B604 CRC64;  
 Query Match 1.1%; Score 10; DB 1; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.36;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0;  
 QY 15 LLLLLLLPL 24  
 Db 19 LLLLLLLPL 28  
 RESULT 13  
 NPW\_RAT STANDARD; PRT; 185 AA.  
 AC O8K1M5;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Neuropeptide W precursor (Preproprotein 18) (pp18) [Contains:  
 DE Neuropeptide W-23 (NPW23) (L8); Neuropeptide W-30 (NPW30) (L8C)].  
 GN NPW OR PP18.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX MEDLINE=22229377; PubMed=12130646;  
 RA Shimomura Y., Harada M., Goto M., Sugo T., Matsumoto Y., Abe M.,  
 RA Watanabe T., Asami T., Kitada C., Mori M., Onda H., Fujino M.;  
 RT "Identification of neuropeptide W as the endogenous ligand for orphan  
 RT G-protein-coupled receptors GPR7 and GPR8.";  
 RL J. Biol. Chem. 277:35826-35832(2002).  
 CC -1- FUNCTION: Plays a regulatory role in the organization of  
 CC neuroendocrine signals accessing the anterior pituitary gland.  
 CC stimulates water drinking and food intake. May play a role in the  
 CC hypothalamic response to stress (By similarity). NPW23 activates  
 CC GPR7 and GPR8 more efficiently than NPW30.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Detected at high levels in the substantia  
 CC nigra, fetal kidney and trachea; at lower levels in testis,  
 CC uterus, ovary and placenta. Not detectable in many regions of the  
 CC central nervous system. Also detected at high levels in  
 CC lymphoblastic leukemia and colorectal adenocarcinoma.  
 CC -1- SIMILARITY: Belongs to the neuropeptide B/W family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL, AB084276; BAC07172.1; ALT\_INIT.  
 KW Neuropeptide; Cleavage on pair of basic residues; Signal.  
 FT SIGNAL 1 32  
 FT PEPTIDE 33 55 NEUROPEPTIDE W-23.  
 FT PROPEP 65 165 NEUROPEPTIDE W-30.  
 FT PROPEP 165 18048 MW; F426AF457F6B604 CRC64;  
 SQ SEQUENCE 165 AA; 18048 MW; F426AF457F6B604 CRC64;  
 Query Match 1.1%; Score 10; DB 1; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.36;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0;  
 QY 15 LLLLLLLPL 24  
 Db 19 LLLLLLLPL 28



Stimulates water drinking and food intake. May play a role in the hypothalamic response to stress. When injected into the lateral cerebromedial, it elevates prolactin (PR) and corticosterone and lowers growth hormone (GH) release.

-1- SUBCELLULAR LOCATION: Secreted.

-1- SIMILARITY: Belongs to the neuropeptide B/w family.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: AB084278; BAC07174.1; -

KW Neuropeptide; Cleavage on pair of basic residues; Signal.

FT SIGNAL 1 41

FT PEPTIDE 42 64 NEUROPEPTIDE W-23.

FT PEPTIDE 42 71 NEUROPEPTIDE W-30.

FT PROPER 74 185 POTENTIAL.

SO SEQUENCE 185 AA; 20231 MW; 2A4C82DFE0A8F6B9 CRC64;

Query Match 1.1%; Score 10; DB 1; Length 185;  
Best Local Similarity 100.0%; Pred. No. 0.4; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

Dr 15 LLLLLLLLLL 24  
28 LLLLLLLLLL 37

RESULT 14  
ATSA BOVIN STANDARD; PRT; 245 AA.

AC Q9PT93; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE ADAMTS-4 (EC 3.4.24.82) (A disintegrin and metalloproteinase with chromospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1) (ADAMP-1) (Fragments).

GN ADAMTS4.

OS Bos taurus (Bovine); Chordata; Craniata; Vertebrata; Euteleostomi; OC Eukaryota; Metazoa; Chordata; Cephalochordata; Ruminantia; Pecora; Bovidae; OC Mammalia; Eutheria; Cetartiodactyla; Kunitzantia; Bovidae; BOV; OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE OF 18-227 FROM N.A.

RX MEDLINE=99333677; PubMed=10403768;

RA Flannery C.R., Little C.B., Hughes C.E., Cateson B.; "Expression of ADAMTS homologues in articular cartilage.";

RT Biochem. Biophys. Res. Commun. 260:318-322(1999).

RL [2]

RN [2]

RP SEQUENCE OF 18-227 FROM N.A.

RX MEDLINE=20092827; PubMed=10625599;

RA Curtis C.L., Hughes C.E., Flannery C.R., Little C.B., Harwood J.L., Cateson B.; "n-3 fatty acids specifically modulate catabolic factors involved in articular cartilage degradation.";

RT J. Biol. Chem. 275:721-724(2000).

RL [3]

RN [3]

RP SEQUENCE OF 1-27; 228-233 AND 234-245.

RX TISSUE=Cartilage; PubMed=10356395;

RX MEDLINE=99286303; PubMed=10356395;

RA Tortorella M.D., Burn T.C., Pratta M.A., Abasrade I., Hollis J.M., Liu R., Rosenfeld S.A., Copeland R.A., Decicco G.P., Wynn R., Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R., Nagase H., Itoh Y., Ellis D.M., Ross H., Wiswall B.H., Murphy K., Hillman M.C., Jr., Hollis G.F., Newton R.C., Magolda R.L., Trzaskos J.M., Arner E.C.; "Purification and cloning of aggrecanase-1: a member of the ADAMTS

family of proteins.";

RL Science 284:1664-1666(1999).

CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be involved in its turnover. May play an important role in the destruction of aggrecan in arthritic diseases. Cleaves aggrecan at the 392-Glu-Ala-393 site.

CC -1- CATALYTIC ACTIVITY: Guanylate endopeptidase; bonds cleaved include 370-Thr-Glu-Gly-Glu-Ala-Arg-Gly-Ser-377 in the interglobular domain of mammalian aggrecan.

CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC -1- INDUCTION: By interleukin-1.

CC -1- PM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC -1- SIMILARITY: Belongs to peptidase family M12B.

CC -1- CAUTION: Has sometimes been referred to as ADAMTS2.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: AF192770; AA07176.1; -

DR HSR: P34179; AATG.

DR InterPro: IPR006025; Pept\_M\_Zn\_BS.

DR InterPro: IPR001590; Peptidase\_M12B.

DR PROSITE: PS00142; ZINC\_PROTEASE; 1.

DR PROSITE: PS50215; ADAM\_MERPO; 1.

KW Hydroxylase; Metalloproteinase; Zinc; Glycoprotein; Extracellular matrix.

FT NON\_TER 1 1

FT NON\_TER 27 28

FT NON\_TER 27 28

FT DOMAIN 27 28

FT METAL 57 58

FT METAL 57 58

FT ACT SITE 58 58

FT METAL 61 61

FT METAL 67 67

FT METAL 227 228

FT NON\_TER 233 234

FT NON\_TER 245 245

FT NON\_TER 245 245

SO SEQUENCE 245 AA; 26391 MW; 7C687968CAB431D CRC64;

Query Match 1.1%; Score 10; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 0.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dr 495 PWADGTPGCP 504  
184 PWADGTPGCP 193

RESULT 15  
CLOC\_HUMAN STANDARD; PRT; 245 AA.

ID CLOC\_HUMAN

AC P02747; Q96DL2; Q96H05;

DT 21-JUL-1986 (Rel. 01, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Complement C1q subcomponent, C chain precursor.

GN C1QG OR C1QC (Human).

OS Homo sapiens (Human); Chordata; Craniata; Vertebrata; Euteleostomi; OC Eukaryota; Metazoa; Chordata; Cephalochordata; Primates; Catarrhini; Hominiidae; Homo.

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX TISSUE=Monocytes; PubMed=1706597;

RX MEDLINE=91174759; PubMed=1706597;

RA Sellar G.C., Blake D.J., Reid K.B.M.; "Characterization and organization of the genes encoding the A-, B-

RT and C-chains of human complement subcomponent C1q. The complete  
RT derived amino acid sequence of human C1q.",  
RL Biochem. J. 274:481-490(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Cerebellum;  
RA Nishi T., Nakagawa S., Senoh A., Mizunuchi H., Inagaki H., Suzuki Y.,  
RA Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,  
RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,  
RA Kawakami B., Nagai K., Isegai T., Sugano S.,  
RA "NEO human CDNA sequencing project";  
RT Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haile F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,  
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Pahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shvachenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez Y.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywanski M.I., Skalska U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 29-122.  
RX MEDLINE=80020137; PubMed=466087;  
RA Reid K.B.M.,  
RT "Complete amino acid sequences of the three collagen-like regions  
RT present in subcomponent C1q of the first component of human  
RT complement";  
RL Biochem. J. 179:367-371(1979).  
RN [5]  
RP REVIEW OF C1Q DEFICIENCY.  
RX MEDLINE=98450587; PubMed=9777412;  
RA Petry F.,  
RT "Molecular basis of hereditary C1q deficiency.";  
RL Immunobiology 199:286-294(1998).  
CC -1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD  
CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE  
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT  
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1  
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE  
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.  
CC -1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R  
CC AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED  
CC OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE  
CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE  
CC C CHAIN.  
CC -1- PTM: O-linked glycans consist of Glc-Gal disaccharides bound to  
CC the oxygen atom of post-translationally added hydroxyl groups.  
CC -1- DISEASE: Defects in C1Q are a cause of C1q deficiency  
CC [MIM:120575]. It is a rare genetic disorder which is associated  
CC with recurrent infections and a high prevalence of lupus  
CC erythematosus-like symptoms. It is characterized by a loss of  
CC activation of the complement classical pathway.  
CC -1- SIMILARITY: Contains 1 C1q domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----  
CC EMBL: AK057792; BAB71575.1; -  
CC EMBL: BC009016; AAC09016.1; -  
CC PIR: S1431; C1HUOC.  
CC GeneW: HGNC:1245; C1QG.  
CC MIM: 120575; -  
CC GO: GO:0005576; C:extracellular; NMS.  
CC GO: GO:0003793; F:defense/immunity protein activity; NMS.  
CC GO: GO:0006955; P:immune response; NMS.  
CC InterPro: IPR001073; C1q.  
CC InterPro: IPR008160; Collagen.  
CC InterPro: IPR008983; TNF-like.  
CC Pfam: PF00386; C1q; 1.  
CC Pfam: PF03391; Collagen; 1.  
CC PRINTS: PR00007; COMPLEMENTC1Q.  
CC SMART: SM0110; C1Q; 1.  
CC PROSITE: PS0113; C1q; 1.  
CC Repeat: signal; Disease mutation.  
CC SIGNAL 1 28  
CC CHAIN 29 245 COMPLEMENT C1Q SUBCOMPONENT, C CHAIN.  
CC DOMAIN 31 112 COLLAGEN-LIKE.  
CC DISULFID 32 32 INTERCHAIN (WITH OTHER C CHAIN).  
CC MOD\_RES 36 36 HYDROXYLATION.  
CC MOD\_RES 39 39 HYDROXYLATION.  
CC MOD\_RES 42 42 HYDROXYLATION.  
CC MOD\_RES 45 45 HYDROXYLATION.  
CC MOD\_RES 54 54 HYDROXYLATION.  
CC MOD\_RES 57 57 HYDROXYLATION.  
CC MOD\_RES 63 63 HYDROXYLATION.  
CC MOD\_RES 66 66 HYDROXYLATION.  
CC MOD\_RES 71 71 HYDROXYLATION.  
CC MOD\_RES 75 75 O-LINKED (GAL. . .).  
CC MOD\_RES 81 81 HYDROXYLATION.  
CC MOD\_RES 84 84 HYDROXYLATION.  
CC MOD\_RES 88 84 O-LINKED (GAL. . .).  
CC MOD\_RES 93 93 HYDROXYLATION.  
CC MOD\_RES 96 96 HYDROXYLATION.  
CC MOD\_RES 99 99 HYDROXYLATION.  
CC MOD\_RES 105 105 HYDROXYLATION.  
CC VARIANT 43 43 G->R (in C1q deficiency).  
CC CONFLICT 14 14 /FTID=VAR 008542.  
CC CONFLICT 23 23 K->R (IN REF. 2).  
CC CONFLICT 27 23 P->A (IN REF. 1).  
CC CONFLICT 57 57 P->P (IN REF. 4).  
CC CONFLICT 66 66 P->K (IN REF. 4).  
CC CONFLICT 72 72 K->P (IN REF. 4).  
CC CONFLICT 84 84 P->K (IN REF. 4).  
CC CONFLICT 87 87 N->D (IN REF. 4).  
CC CONFLICT 90 90 M->N (IN REF. 4).  
CC CONFLICT 215 215 E->G (IN REF. 2).  
CC SEQUENCE 245 AA; 25774 MW; FAL7117EB7ABFC12 CRC64;  
Query Match 1.1%; Score 10; DB 1; Length 245;  
Best local similarity 100.0%; Pred No. 0.5; Indels 0; Gaps 0;  
Matches 10; Mismatches 0;  
QY 15 LLLLLLLPL 24  
DB 15 LLLLLLLPL 24

Search completed: May 7, 2004, 11:59:06  
Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:42:50 ; Search time 65 Seconds  
(without alignments)  
3668.727 Million cell updates/sec

Title: US-09-989-687-4  
Perfect score: 4853  
Sequence: 1 MFPAAPAPRWLPFLLLLL.....CNKAKPEDAKPCESQLCPL 890

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp19808:\*\n2: Geneseqp19908:\*\n3: Geneseqp20008:\*\n4: Geneseqp20018:\*\n5: Geneseqp20028:\*\n6: Geneseqp20038:\*\n7: Geneseqp20038:\*\n8: Geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4853	100.0	890	2	AAV49502 Human MET
2	4853	100.0	890	2	AAV49502 Human MET
3	4822	99.4	890	6	ABP96306 Human ADA
4	4807.5	99.1	889	4	AAV74946 Human ADA
5	3919.5	60.8	905	4	AAV72284 Human ADA
6	3713	76.5	680	3	AAV21251 Human met
7	3244	48.3	896	3	AAV21265 Human met
8	2344	48.3	896	6	ABU08387 Murine ma
9	2328.5	48.0	950	4	AAV49501 Human MET
10	2328.5	48.0	950	4	AAV73549 Human ADA
11	2328.5	48.0	950	4	AAV50002 Human MET
12	2328.5	48.0	950	2	AAV80285 Human int
13	2328.5	48.0	968	4	AAV50011 Protein;
14	2327.5	48.0	967	2	AAV04142 Human Tan
15	2317.5	47.8	950	3	AAV53899 Human act
16	2315.5	47.7	967	2	AAV78189 Human sec
17	2315.5	47.7	967	6	AAV57139 Human sec
18	2315.5	47.7	967	6	AAV41003 Human sec
19	2315.5	47.7	967	6	AAV91631 Human sec
20	2315.5	47.7	967	7	AAV74267 Human sec
21	2315.5	47.7	967	7	AAV73948 Human sec
22	2280.5	47.0	481	2	AAV04145 Rat Tango
23	2146.5	44.2	924	5	ABP70062 Human NOV
24	2138.5	44.1	950	5	AAV22541 Human pro
25	2138.5	44.1	950	6	ABU09520 Human pro

26	2135.5	44.0	950	4	AAV62299 Human met
27	2131.5	43.9	950	6	ABR40092 Human ADA
28	2115.5	43.6	952	5	AAV74751 Human ADA
29	2103.5	43.3	727	2	AAV78435 Human ADA
30	2043.5	42.1	928	5	AAV72899 Human met
31	1985.5	41.0	823	6	ABU08383 Human met
32	1989	41.0	364	4	AAV09710 Human gen
33	1989	41.0	364	7	AAV22088 Human pro
34	1842.5	38.0	837	2	AAV75425 Human ag9
35	1842.5	38.0	837	4	AAV78228 Human ag9
36	1842.5	38.0	837	7	AAV85488 Human ag9
37	1842.5	38.0	840	3	AAV21256 Human met
38	1841.5	37.9	837	3	AAV99429 Human pro
39	1841.5	37.9	837	4	AAV66178 Human pro
40	1841.5	37.9	837	4	AAV29199 Human pro
41	1841.5	37.9	837	6	ABU58575 Human pro
42	1841.5	37.9	837	6	ABU88123 Human pro
43	1841.5	37.9	837	6	ABU84438 Human sec
44	1841.5	37.9	837	6	ABR66312 Human sec
45	1841.5	37.9	837	6	ABR65702 Human sec

## ALIGNMENTS

RESULT 1  
AAV49502  
ID AAV49502 standard; protein; 890 AA.

XX AC AAV49502;  
XX DT 10-JAN-2000 (first entry)  
XX DE Human METH2 protein.  
XX DE Human METH2 protein.  
XX KW Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;  
XX KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;  
XX KW angiogenesis inhibitor; abnormal wound healing; inflammation;  
XX KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;  
XX KW diabetic retinopathy; macula degeneration; haemangioma; detection;  
XX KW arterial-venous malformation; immune deficiency.  
XX OS Homo sapiens.  
XX PN WO9937660-A1.  
XX PD 29-JUL-1999.  
XX PF 22-JAN-1999; 99WO-US001313.  
XX PR 23-JAN-1998; 98US-0072298P.  
XX PR 28-AUG-1998; 98US-0098539P.  
XX PA (IRU//) IRU8A-Arispe L.  
XX PA (HAST//) HASTINGS G A.  
XX PA (ROBE//) ROBEN S W.  
XX PI Irrela-Arispe L, Hastings GA, Ruben SM;  
XX DR WPI, 1999-590684/50.  
XX DR N-PSDB; AA232001.  
XX PT New isolated metalloprotease thrombospondin polypeptides, useful for  
XX PT treating hyperproliferative disorders, cancers or autoimmune disorders.  
XX PS Claim 10; Fig 2; 457pp; English.  
XX CC AA232000 and AA232001 encode, and AAV49501 and AAV49502 represent, human  
XX CC metalloprotease thrombospondin (METH) proteins METH1 and METH2  
XX CC respectively. METH1 and METH2 have been found to be potent inhibitors of  
XX CC angiogenesis both in vitro and in vivo. They can be used for treating  
XX CC cancer and other disorders related to angiogenesis including abnormal  
XX CC wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial

bleeding disorders, diabetic retinopathy, some forms of macula degeneration, haemangiomas, and arterial-venous malformations. They may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilisation (chemotaxis) of immune cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or infectious. They can also be used to treat inflammatory conditions, both chronic and acute conditions. The products can also be used for detection and diagnosis. AA32002 to AA32080, and AA39503 to AA49511 represent sequences given in the exemplification of the present invention

Sequence 890 AA;

Query Match 100.0%; Score 4853; DB 2; Length 890;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPAAPAPRMPLPILLLILLPLARAPAPAPAGGASATVPTLPAGAGLALHLSA 60  
DB 1 MFPAAPAPRMPLPILLLILLPLARAPAPAPAGGASATVPTLPAGAGLALHLSA 60  
QY 61 FGKGFVRLAPDDSFLLAPEFKIERLGSGGATGGERGLRCGCFSGTVNGEPESLAAVSLC 120  
DB 61 FGKGFVRLAPDDSFLLAPEFKIERLGSGGATGGERGLRCGCFSGTVNGEPESLAAVSLC 120  
QY 121 RGLSGSFLLDGEFTTPOGAGSGLAOPHRLQRMGPAPAPLRGPEWEVETEGEQORER 180  
DB 121 RGLSGSFLLDGEFTTPOGAGSGLAOPHRLQRMGPAPAPLRGPEWEVETEGEQORER 180  
QY 181 GDHGEDSEESQEEAEAGASPPPLGTSRTKAFSEAEFVEFLVAADASMAAFYCADL 240  
DB 181 GDHGEDSEESQEEAEAGASPPPLGTSRTKAFSEAEFVEFLVAADASMAAFYCADL 240  
QY 181 GDHGEDSEESQEEAEAGASPPPLGTSRTKAFSEAEFVEFLVAADASMAAFYCADL 240  
DB 181 GDHGEDSEESQEEAEAGASPPPLGTSRTKAFSEAEFVEFLVAADASMAAFYCADL 240  
QY 241 QNHILITMSVAAIYKHSINSINLMVVKYLVEDEKMGPEVSDNGLLTRNFCWQORR 300  
DB 241 QNHILITMSVAAIYKHSINSINLMVVKYLVEDEKMGPEVSDNGLLTRNFCWQORR 300  
QY 301 FNQPSDRPREHYTALILTRQNFQGEGLCDTLGAVDITGCDPNKSCSYIEBEGLOAH 360  
DB 301 FNQPSDRPREHYTALILTRQNFQGEGLCDTLGAVDITGCDPNKSCSYIEBEGLOAH 360  
QY 301 FNQPSDRPREHYTALILTRQNFQGEGLCDTLGAVDITGCDPNKSCSYIEBEGLOAH 360  
DB 301 FNQPSDRPREHYTALILTRQNFQGEGLCDTLGAVDITGCDPNKSCSYIEBEGLOAH 360  
QY 361 TLAEHLGVLSPHDSKPCRTLFGMGKHVMAPLFVHNLQTLPMSPCSAMYTELLDG 420  
DB 361 TLAEHLGVLSPHDSKPCRTLFGMGKHVMAPLFVHNLQTLPMSPCSAMYTELLDG 420  
QY 421 GHGDCILAPGAPALPLPTGLGPRMALYOLDQCCQIIEBDRPHCPNTSAOVCAQLMCHT 480  
DB 421 GHGDCILAPGAPALPLPTGLGPRMALYOLDQCCQIIEBDRPHCPNTSAOVCAQLMCHT 480  
QY 481 DGAEPILCTKNGSLPMADGTCGPGHLCSESGCLPEEEVERPKRYVDGMAPWMPWEGCS 540  
DB 481 DGAEPILCTKNGSLPMADGTCGPGHLCSESGCLPEEEVERPKRYVDGMAPWMPWEGCS 540  
QY 541 RTGGGVOFSHRECDPEPONGRCYCLGRAPKQSCHEBCECPDEKSKRECKINAYN 600  
DB 541 RTGGGVOFSHRECDPEPONGRCYCLGRAPKQSCHEBCECPDEKSKRECKINAYN 600  
QY 601 YTDMDGNLLOWPKYAGVSPRDRCLFCRARGRSEFFKFEAKVIDGLTGEETLACVRG 660  
DB 601 YTDMDGNLLOWPKYAGVSPRDRCLFCRARGRSEFFKFEAKVIDGLTGEETLACVRG 660  
QY 661 QCVKAGCDHVDSPRLDKGVCYGKNGSKRVSSGLPTNYGNDYITIPAGANTIDVK 720  
DB 661 QCVKAGCDHVDSPRLDKGVCYGKNGSKRVSSGLPTNYGNDYITIPAGANTIDVK 720  
QY 721 QSHSPVONDNVYALKTADGQYLLNGNLAISATGCDILVKGTILKYGSSATATERLOS 780  
DB 721 QSHSPVONDNVYALKTADGQYLLNGNLAISATGCDILVKGTILKYGSSATATERLOS 780  
QY 781 RPLPEPLVQOLLTPGEVEPPKYKTYFFVNDVDSMOSSKRRATNTIIOPLHAWMLG 840  
DB 781 RPLPEPLVQOLLTPGEVEPPKYKTYFFVNDVDSMOSSKRRATNTIIOPLHAWMLG 840

QY 841 DWSGCSSTCGAGWQRRVTEGCRDPGGSATCNKALKREDAKPCESQICPL 890  
DB 841 DWSGCSSTCGAGWQRRVTEGCRDPGGSATCNKALKREDAKPCESQICPL 890

# RESULT 2

AA350003 standard; protein; 890 AA.

AA350003;

19-MAR-2001 (first entry)

Human METH2.

Human; METH2; metalloprotease; thrombospondin; angiogenesis inhibition; cancer therapy; benign tumor; ocular angiogenic disease; rheumatoid arthritis; psoriasis; wound healing; endometriosis; vasculogenesis; granulation; hypertrophic scar; nonunion fracture; scleroderma; trachoma; vascular adhesion; myocardial angiogenesis; coronary collateral; cerebral collateral; arteriovenous malformation; coronary limb angiogenesis; Osler-Webber syndrome; ischaemic limb angiogenesis; telangiectasia; haemophilic joint; plaque neovascularisation; birth control; Crohn's disease; atherosclerosis; birth control.

Hom sapiens.

WO200071577-A1.

30-NOV-2000.

25-MAY-2000; 2000WO-US014462.

25-MAY-1999; 99US-00318208.  
20-DEC-1999; 99US-0144882P.  
10-AUG-1999; 99US-0147822P.  
13-AUG-1999; 99US-00373658.  
22-DEC-1999; 99US-0171503P.  
22-FEB-2000; 2000US-0183792P.

(HUMA-) HUMAN GENOME SCI INC.

(SMK) SMITHKLINE BEECHAM CORP.

(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

(IRUE) IRUELA-ARISPE L.

(HAST) HASTINGS G A.

(RUBE) RUBEN S M.

(JONK) JONAK Z L.

(TRUL) TRULLI S H.

(FORN) FORNWALD J A.

(TERR) TERRRETT J A.

IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH; Fornwald JA, Terrrett JA;

WPI; 2001-025136/03.

N-PSDB; AAC90058.

METH1 and METH2 polynucleotides and encoded polypeptides, used to inhibit angiogenesis in the treatment of disorders such as cancer, rheumatoid arthritis and psoriasis.

Claim 15; Fig 2; 768pp; English.

The present sequence is human METH2 (ME for metalloprotease and TH for thrombospondin). METH2 can be used for inhibiting angiogenesis in an individual, and for treating cancer, benign tumors, an ocular angiogenic disease, rheumatoid arthritis, psoriasis, delayed wound healing, nonunion fracture, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque

CC neovascularisation, telangiectasia, haemophiliac joints, angiodysplasia,  
 CC fibromuscular dysplasia, wound granulation, Crohn's disease or  
 CC atherosclerosis. METH2 can also be used in bith control. METH2 can also  
 CC be used in diagnostic methods for the prognosis of cancer

XX  
 XX  
 SQ Sequence 890 AA:

Query Match 100.0%; Score 4853; DB 4; Length 890;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPAAPARWLPFLILLILLPLARGAPAPAPAGGASLVVPTPLPSAGELALHLSA 60  
 DB 1 MFPAAPARWLPFLILLILLPLARGAPAPAPAGGASLVVPTPLPSAGELALHLSA 60  
 QY 61 FGKGFVRLAPDSDFLAPEFKIERLGSGGRATGGERLRCGCFSGTVNGEPBSLAVSIC 120  
 DB 61 FGKGFVRLAPDSDFLAPEFKIERLGSGGRATGGERLRCGCFSGTVNGEPBSLAVSIC 120  
 QY 121 RGLSGSFLDGESEFTIQPGAGGSLAQPHRLQRMGPAGAPLPRGPEWEVETGEGROER 180  
 DB 121 RGLSGSFLDGESEFTIQPGAGGSLAQPHRLQRMGPAGAPLPRGPEWEVETGEGROER 180  
 QY 181 GDHODESEESQEEAEAGASPPPLGATSRTRFVSEARFVETLLVADASMAAFYADL 240  
 DB 181 GDHODESEESQEEAEAGASPPPLGATSRTRFVSEARFVETLLVADASMAAFYADL 240  
 QY 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLIYEDKMGPEVSDNGGLTLRNCWQR 300  
 DB 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLIYEDKMGPEVSDNGGLTLRNCWQR 300  
 QY 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLIYEDKMGPEVSDNGGLTLRNCWQR 300  
 DB 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLIYEDKMGPEVSDNGGLTLRNCWQR 300  
 QY 301 FNOPSDRPHRYDAILLTRONFCGEGLCDTLGVADIGTICPNNSCSVTEBEGIAAH 360  
 DB 301 FNOPSDRPHRYDAILLTRONFCGEGLCDTLGVADIGTICPNNSCSVTEBEGIAAH 360  
 QY 361 TLAHELGVLSMPHDSKPCRTLFGPMGKHVMAFLFVHLNQLTLPMSGSAITLTDG 420  
 DB 361 TLAHELGVLSMPHDSKPCRTLFGPMGKHVMAFLFVHLNQLTLPMSGSAITLTDG 420  
 QY 421 GHGQCLDAPGAPALPLPTGLPGRNALYQLDQCQQLGPFRCFNTSADVCAQLMCHT 480  
 DB 421 GHGQCLDAPGAPALPLPTGLPGRNALYQLDQCQQLGPFRCFNTSADVCAQLMCHT 480  
 QY 481 DGAEPLCHTKNGSLPMADGTPCGPGLCSGSCLPSEEVPRPRVVDGAPMGWEGCS 540  
 DB 481 DGAEPLCHTKNGSLPMADGTPCGPGLCSGSCLPSEEVPRPRVVDGAPMGWEGCS 540  
 QY 541 RTGGGVOFHSRECKDPONGGRVCLGRAXOSCHTEECPPDGKSPRGOCEKNAYN 600  
 DB 541 RTGGGVOFHSRECKDPONGGRVCLGRAXOSCHTEECPPDGKSPRGOCEKNAYN 600  
 QY 601 YTTMDGNLLQWVPKXAGVSPDRCKLFCRARGSEFVFAKYIDGTLGCEPITALCVAG 660  
 DB 601 YTTMDGNLLQWVPKXAGVSPDRCKLFCRARGSEFVFAKYIDGTLGCEPITALCVAG 660  
 QY 661 YTTMDGNLLQWVPKXAGVSPDRCKLFCRARGSEFVFAKYIDGTLGCEPITALCVAG 660  
 DB 661 YTTMDGNLLQWVPKXAGVSPDRCKLFCRARGSEFVFAKYIDGTLGCEPITALCVAG 660  
 QY 661 OCYVAGCHVVDSPKDKCGVCGGKNSCKXVSGSLTPNNGNDIVITPAATIDVX 720  
 DB 661 OCYVAGCHVVDSPKDKCGVCGGKNSCKXVSGSLTPNNGNDIVITPAATIDVX 720  
 QY 721 QRSHPGVNDGNVYALKTADQVYLLNGNLAISAEODILVKGTLKXSGSIATLERLOS 780  
 DB 721 QRSHPGVNDGNVYALKTADQVYLLNGNLAISAEODILVKGTLKXSGSIATLERLOS 780  
 QY 781 RPLPEPLTVOLLTPGVFPKVKYTFVPVNDVPSOSSKERATNIIOTPLHAWMG 840  
 DB 781 RPLPEPLTVOLLTPGVFPKVKYTFVPVNDVPSOSSKERATNIIOTPLHAWMG 840  
 QY 841 DMSECSSTCGAGWQRTVECRDPGQASATCNKALKPEDAKPCESQICPL 890  
 DB 841 DMSECSSTCGAGWQRTVECRDPGQASATCNKALKPEDAKPCESQICPL 890

RESULT 3

ABP96306  
 ID ABP96306 standard; protein: 890 AA.

XX AC ABP96306;

XX 20-MAY-2003 (first entry)

XX Human ADAMTS8 protein.

XX Humanised baculovirus; cytosstatic; gene therapy; baculovirus; cancer;  
 XX prostate cancer; chromosome 11.

XX Homo sapiens.

XX MO2003016540-A2.

XX 27-FEB-2003.

XX 15-AUG-2002; 2002WO-GB003791.

XX 15-AUG-2001; 2001GB-00019852.

XX (UYVO-) UNIV YORK.

XX Maitland N;

XX WPI; 2003-268336/26.

PT New baculovirus having a modified genome encoding a therapeutic agent,  
 PT useful in the manufacture of a medicament for the treatment of cancer,  
 PT particularly prostate cancer.

XX Claim 24; Page; 34pp; English.

CC The present invention describes a humanised baculovirus (1) which  
 CC comprises a modified baculovirus genome having a nucleic acid molecule  
 CC encoding a therapeutic agent and a polypeptide which functions to target  
 CC the baculovirus to at least one cell type. Also described is a  
 CC pharmaceutical composition comprising (1). (1) has cytostatic activity,  
 CC and can be used in gene therapy. The baculovirus is useful in the  
 CC manufacture of a medicament for the treatment of cancer, particularly  
 CC prostate cancer. The present sequence represents the human disintegrin-  
 CC like and metalloprotease (teprolysin type) with thrombospondin type 1  
 CC motif, 8 (ADMTS8) protein, which is specified in the exemplification of  
 CC the present invention as angiotensin. N.B. The present sequence is not  
 CC given in the specification but is referred to in Claim 24 as Genbank  
 CC accession number NM\_007037

XX Sequence 890 AA:

XX Query Match 99.4%; Score 4822; DB 6; Length 890;

XX Best Local Similarity 99.4%; Pred. No. 0;

XX Matches 885; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MFPAAPARWLPFLILLILLPLARGAPAPAPAGGASLVVPTPLPSAGELALHLSA 60  
 DB 1 MFPAAPARWLPFLILLILLPLARGAPAPAPAGGASLVVPTPLPSAGELALHLSA 60  
 QY 61 FGKGFVRLAPDSDFLAPEFKIERLGSGGRATGGERLRCGCFSGTVNGEPBSLAVSIC 120  
 DB 61 FGKGFVRLAPDSDFLAPEFKIERLGSGGRATGGERLRCGCFSGTVNGEPBSLAVSIC 120  
 QY 121 RGLSGSFLDGESEFTIQPGAGGSLAQPHRLQRMGPAGAPLPRGPEWEVETGEGROER 180  
 DB 121 RGLSGSFLDGESEFTIQPGAGGSLAQPHRLQRMGPAGAPLPRGPEWEVETGEGROER 180  
 QY 181 GDHODESEESQEEAEAGASPPPLGATSRTRFVSEARFVETLLVADASMAAFYADL 240  
 DB 181 GDHODESEESQEEAEAGASPPPLGATSRTRFVSEARFVETLLVADASMAAFYADL 240  
 QY 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLIYEDKMGPEVSDNGGLTLRNCWQR 300  
 DB 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLIYEDKMGPEVSDNGGLTLRNCWQR 300

```

QY 301 FNPSPDRPHRYDTAILLTRQNFQGGEGLCDTLGVADIGTICDPNKSQSVIEDGLQAH 360
DB 301 FNPSPDRPHRYDTAILLTRQNFQGGEGLCDTLGVADIGTICDPNKSQSVIEDGLQAH 360
QY 361 TLNHELGHVLSMHDSPKCTRLFGPMGKHVMAPLFVHLNQTLPMSPCSAWLTLELDG 420
DB 361 TLNHELGHVLSMHDSPKCTRLFGPMGKHVMAPLFVHLNQTLPMSPCSAWLTLELDG 420
QY 421 GHGDCILDAPGAALPLPTGLPGEMALYOLDQCRQIFGPDFRHCNPTSAQDVCAQIMCHT 480
DB 421 GHGDCILDAPRALPLPTGLPGEMALYOLDQCRQIFGPDFRHCNPTSAQDVCAQIMCHT 480
QY 481 DGAEPILCHTKNGSLPMADGTPCGPHLCSBGLPEEVEVERPKPVADGMAFMGMBGCS 540
DB 481 DGAEPILCHTKNGSLPMADGTPCGPHLCSBGLPEEVEVERPKPVADGMAFMGMBGCS 540
QY 541 RTGGGVQFSHRECKDPEPONGRKYCLGRBAKYQSCHTBECPPDGKSPREOQCEKYNAYN 600
DB 541 RTGGGVQFSHRECKDPEPONGRKYCLGRBAKYQSCHTBECPPDGKSPREOQCEKYNAYN 600
QY 601 YTTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFFVFAKYIDGTLGSPETLAI CVRG 660
DB 601 YTTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFFVFAKYIDGTLGSPETLAI CVRG 660
QY 661 QCVKAGCDHVDSFRKLDKCGVCGKGNSCRKVSGLPTPNYGYNDIVITIPAGATNIDYK 720
DB 661 QCVKAGCDHVDSFRKLDKCGVCGKGNSCRKVSGLPTPNYGYNDIVITIPAGATNIDYK 720
QY 721 QRSHPGVQNDGNVYALKTADGQYLLNGNLAI SAIBODILVKGTILKXSGSIATLERLOSF 780
DB 721 QRSHPGVQNDGNVYALKTADGQYLLNGNLAI SAIBODILVKGTILKXSGSIATLERLOSF 780
QY 781 RPLPEPLTVQLLTVPGEVFPKPYKTFEVPNDVDFSMQSSKERATTNIIQPLIHAQWVIG 840
DB 781 RPLPEPLTVQLLTVPGEVFPKPYKTFEVPNDVDFSMQSSKERATTNIIQPLIHAQWVIG 840
QY 841 DMSBSSSTCGAGWQRTVECRDPSSGASATCNKALKEPAKPCESQICPL 890
DB 841 DMSBSSSTCGAGWQRTVECRDPSSGASATCNKALKEPAKPCESQICPL 890

```

RESULT 4  
AAB74946  
ID AAB74946 standard; protein; 889 AA.  
AC AAB74946;  
DT 02-JUL-2001 (first entry)  
DE Human ADAM type metal protease MDT3 protein SEQ ID NO:20.  
XX Human; a disintegrin and metalloprotease type metal protease; MDT3;  
XX MDT3; ADAM type metal protease; cytosolic; antiarthritic;  
XX cancer; arthritis; arthrosis deformans.  
OS Homo sapiens.  
XX  
XX JP2001008687-A.  
XX  
XX 16-JAN-2001.  
XX  
XX 25-JUN-1999; 99JP-00180973.  
XX  
XX 25-JUN-1999; 99JP-00180973.  
XX  
XX (YAMA) YAMANOUCHI PHARM CO LTD.  
XX  
XX WPI: 2001-285362/30.  
XX  
XX N-PSDB; AAF82166.  
XX  
XX New metal protease and metal protease gene, for use as a drug for  
XX cancers, arthritis and arthrosis deformans.  
PT

```

XX Claim 1; Page 25-27; 31pp; Japanese.
BS The present sequence represents a disintegrin and metalloprotease (ADAM)
XX type metal protease designated MDT3, isolated from human. MDT3 proteins
CC have cytosolic and antiarthritic activities. They can be used as a drug
CC for cancers, arthritis and arthrosis deformans
XX
XX Sequence 889 AA:
Query Match 99.1%; Score 4807.5; DB 4; Length 889;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 884; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

```

1 MFPAAPRMLPFLILLILLPLARGAAPAPAAQAGASLIVPTRLPGSAGELALHLSA 60  
1 MLPAAPRMPF-LILLILLPLARGAAPAPAAQAGASLIVPTRLPGSAGELALHLSA 59  
61 FGKGFVRLADDSFLAPDEFKIERLGGSGATGGERLRCGFPGCTVNGPESLAAVSLC 120  
60 FGKGFVRLADDSFLAPDEFKIERLGGSGATGGERLRCGFPGCTVNGPESLAAVSLC 119  
121 RGLSGFLDDEEFTIQOGAGSLAOPHRLQRMWPAARPLRGPMEVETGEGQOER 180  
120 RGLSGFLDDEEFTIQOGAGSLAOPHRLQRMWPAARPLRGPMEVETGEGQOER 179  
181 GDHQDSEESQEEBAGASPPPLGATSTTKPVSEARFVETLIVADASMAAFYQADL 240  
180 GDHQDSEESQEEBAGASPPPLGATSTTKPVSEARFVETLIVADASMAAFYQADL 239  
241 QNHILTLMSVARIYKPSIKNSINLMVVKYLIVDEKMGVSDNGSLTLRNCNMQRR 300  
240 QNHILTLMSVARIYKPSIKNSINLMVVKYLIVDEKMGVSDNGSLTLRNCNMQRR 299  
301 FNPSPDRPHRYDTAILLTRQNFQGGEGLCDTLGVADIGTICDPNKSQSVIEDGLQAH 360  
300 FNPSPDRPHRYDTAILLTRQNFQGGEGLCDTLGVADIGTICDPNKSQSVIEDGLQAH 359  
361 TLNHELGHVLSMHDSPKCTRLFGPMGKHVMAPLFVHLNQTLPMSPCSAWLTLELDG 420  
360 TLNHELGHVLSMHDSPKCTRLFGPMGKHVMAPLFVHLNQTLPMSPCSAWLTLELDG 419  
421 GHGDCILDAPGAALPLPTGLPGEMALYOLDQCRQIFGPDFRHCNPTSAQDVCAQIMCHT 480  
420 GHGDCILDAPRALPLPTGLPGEMALYOLDQCRQIFGPDFRHCNPTSAQDVCAQIMCHT 479  
481 DGAEPILCHTKNGSLPMADGTPCGPHLCSBGLPEEVEVERPKPVADGMAFMGMBGCS 540  
480 DGAEPILCHTKNGSLPMADGTPCGPHLCSBGLPEEVEVERPKPVADGMAFMGMBGCS 539  
541 RTGGGVQFSHRECKDPEPONGRKYCLGRBAKYQSCHTBECPPDGKSPREOQCEKYNAYN 600  
540 RTGGGVQFSHRECKDPEPONGRKYCLGRBAKYQSCHTBECPPDGKSPREOQCEKYNAYN 599  
601 YTTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFFVFAKYIDGTLGSPETLAI CVRG 660  
600 YTTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFFVFAKYIDGTLGSPETLAI CVRG 659  
661 QCVKAGCDHVDSFRKLDKCGVCGKGNSCRKVSGLPTPNYGYNDIVITIPAGATNIDYK 720  
660 QCVKAGCDHVDSFRKLDKCGVCGKGNSCRKVSGLPTPNYGYNDIVITIPAGATNIDYK 719  
721 QRSHPGVQNDGNVYALKTADGQYLLNGNLAI SAIBODILVKGTILKXSGSIATLERLOSF 780  
720 QRSHPGVQNDGNVYALKTADGQYLLNGNLAI SAIBODILVKGTILKXSGSIATLERLOSF 779  
781 RPLPEPLTVQLLTVPGEVFPKPYKTFEVPNDVDFSMQSSKERATTNIIQPLIHAQWVIG 840  
780 RPLPEPLTVQLLTVPGEVFPKPYKTFEVPNDVDFSMQSSKERATTNIIQPLIHAQWVIG 839  
841 DMSBSSSTCGAGWQRTVECRDPSSGASATCNKALKEPAKPCESQICPL 890  
840 DMSBSSSTCGAGWQRTVECRDPSSGASATCNKALKEPAKPCESQICPL 889

```

RESULT 5
AAB72284
ID AAB72284 standard; protein; 905 AA.
XX
AC AAB72284;
XX
DT 14-MAY-2001 (first entry)
XX
DE Murine ADAMTS-8 amino acid sequence.
XX
KW ADAMTS-N; disintegrin; metalloprotease; thrombospondin type I motif;
KW tumour cachexia; inflammation; dermatosparaxis; EDS-VIIC; angiogenesis;
KW Ehlers-Danlos syndrome type VIIC; articular cartilage erosion; mouse;
KW metastasis; embryogenesis; egg implantation; ADAMTS-8.
XX
OS Mus musculus.
XX
PN M020011074-A2.
XX
PD 15-FEB-2001.
XX
PF 03-AUG-2000; 2000MO-US021223.
XX
PR 06-AUG-1999; 99US-00369364.
XX
PA (CLEV-) CLEYELAND CLINIC FOUND.
PA (APTE-) APTE S. S.
PA (HURS-) HURSKAINEN T L.
PA (HIRO-) HIROHATA S.
XX
PI Apte SS, Hurskainen TL, Hirohata S;
XX
DR WPI, 2001-159978/16.
XX
DR N-PSDB; AAF63441.
XX
PT Murine and human 'A Disintegrin-like And Metalloprotease domain with
PT Thrombospondin type I motifs' proteins and the nucleic acids encoding
PT them, useful for treating e.g. tumors, inflammation and arthritis.
XX
PS Claim 1; Fig 5; 181pp; English.
XX
CC This invention relates to murine and human ADAMTS-N (A disintegrin-like
CC and metalloprotease domain with thrombospondin type I motifs) proteins,
CC designated ADAMTS-5, 6, 7, 8, 9, 10 and R1. Also included in the
CC invention are cDNA sequences encoding the proteins, and antibodies
CC specific for the proteins. The nucleic acid sequences and proteins may be
CC used in the prevention, diagnosis and treatment of diseases associated
CC with inappropriate ADAMTS-N expression. Disorders that may be treated
CC using the nucleic acids, proteins and antibodies include, for example
CC tumour cachexia, inflammation, dermatosparaxis in cattle or Ehlers-Danlos
CC syndrome type VIIC (EDS-VIIC) in humans, erosion of articular cartilage
CC in arthritic (both inflammatory and non-inflammatory) disease,
CC angiogenesis, tumour growth and metastases, and they may also be used for
CC controlling embryogenesis and implantation of fertilised eggs. The
CC present sequence represents murine ADAMTS-8
XX
SQ Sequence 905 AA;
XX
Query Match 80.8%; Score 3919.5; DB 4; Length 905;
Best Local Similarity 79.2%; Pred. No. 3e-269;
Matches 717; Conservative 74; Mismatches 99; Indels 15; Gaps 4;
XX
QY 1 MFPAAPARWLPFLILLILL-LPLARCAPAPPAAGQASLIVPTRLPAGSLAHL 59
DB 1 MLRDPITGWPILLILLQLPPEPPVVCAPAPGPGTGAQASLIVPTRLPAGSLAHL 60
QY 60 AFKGFVRLAPDDSFILAPFKILRGSGRATGGERGLRGCFPSGTVNGSPESLA 119
DB 61 AFGGQFVRLAPDDSFILAPFKILRGSGRATGGERGLRGCFPSGTVNGRSLA 120
QY 120 CRGLSGFLLDGEFTIQGAGGSLAOPHRLQRWGAPAR-----PLPRGP 168

```

```

DB 121 VAGWSGFLLAGBEFTIQPGAGDSLDQPHRLQRWGPGQGRREDPGLAAAEVFLPQGLEW 180
QY 169 EVETEGEQROERGDHODESEBGOE--EAEAGASEPPPLGATSRTRFVSEARFVETLL 226
DB 181 EEWGNQGGQGRSDNEDEKQDXEGLKTEDESKRPVPPGSKTRSRVSEARFVETLL 240
QY 227 VADASMAAFYADLQNHILTLMSVAARIYKPSIKNSINLMVVKVILVEDEKMGPEVSDN 286
DB 241 VADASMAAFYCTDQNHILTLMSVAARIYKPSIRNSVNLVYVYVILVEKERMPEVSDN 300
QY 287 GGLTLRNFQWRRFENQPSDRHPHYDTAILLIRONCGQEGCLDTLGVADIGTCDPNK 346
DB 301 GGLTLRNFQWRRFENQPSDRHPHYDTAILLIRONCGQEGCLDTLGVADIGTCDPNK 360
QY 347 SCSTVEEGLOAATLHLHELGHVLSMHPDQSKPTRLFQPMGKHVNAPLFVHINQTLFW 406
DB 361 SCSTVEEGLOAATLHLHELGHVLSMHPDQSKPTRLFQPMGKHVNAPLFVHINQTLFW 420
QY 407 SPCSAVYLTTELDDGHDCLLDAPGAALPLPTGLPGRMALYOLDQCRQIFGDPFRICPN 466
DB 421 SPCSAVYLTTELDDGHDCLLDAPGAALPLPTGLPGRMALYOLDQCRQIFGDPFRICPN 480
QY 467 TSAQDVCAQLWC-HTDAEPLCHTKNGSLPMDGTFCGPGHLCSEGSCLPEEYERPKV 525
DB 491 TSVEDICVQLCARHRDSEPLCHTKNGSLMADGTPCGPHLCIDGSCVLKEDVENPKAV 540
QY 526 VDGGWAPWGPWGBSCRTGGGVQFSHECKDPEPONGGRVYCLGRARYQSCHTBECPPDG 585
DB 541 VDGGWAPWGPWGBSCRTGGGVQFSHECKDPEPONGGRVYCLGRARYQSCHTBECPPDG 600
QY 586 KSFRQCEKKNAYNTMDNLLQWPKYGVSPEDRCXLFCAARGSEKVEAKYID 645
DB 601 KSFRQCEKKNAYNTMDNLLQWPKYGVSPEDRCXLFCAARGSEKVEAKYID 660
QY 646 GTLCGPETLAIQVAGQCVKAGCDHYVDSPRKLDKCGVCGKNSCRKXVSGSLPTNNGYN 705
DB 661 GTLCGPETLAIQVAGQCVKAGCDHYVDSPRKLDKCGVCGKNSCRKXVSGSLPTNNGYN 720
QY 706 DIVTPAGATNIDVQSRHPGVQNDGNTLAKTKDQGLTNGNLASIEEDILVKGITL 765
DB 721 DIVTPAGATNIDVQSRHPGVQNDGNTLAKTKDQGLTNGNLASIEEDILVKGITL 780
QY 766 KYSGSIATLERLQSRPPEPLTVQLTVPEVEPPKRYTFEVPNDVDFSMQSKERAT 825
DB 781 KYSGSIATLERLQSRPPEPLTVQLTVPEVEPPKRYTFEVPNDVDFSMQSKERAT 840
QY 826 TNIQPLHLAOWVLGDWSECSSTCGAGQRTVECRDPSGQASATCNKALPREDAKPES 885
DB 841 TNIQPLHLAOWVLGDWSECSSTCGAGQRTVECRDPSGQASATCNKALPREDAKPES 900
QY 886 QICPL 890
DB 901 QICPL 905

```

```

RESULT 6
AAB21251
ID AAB21251 standard; protein; 680 AA.
XX
AC AAB21251;
XX
DT 23-FEB-2001 (first entry)
XX
DE Human metalloproteinase ADAMTS-2.
XX
KW Human; ADAMTS2; metalloproteinase; ADAM;
KW a disintegrin and metalloproteinase domain; thrombospondin domain;
KW vaccine; nontoxic; neuroprotective; antiparkinsonian; cerebroprotective;
KW cytostatic; antiarthritic; immunosuppressive; Alzheimer's disease;
KW Parkinson's disease; stroke; cancer; arthritis; autoimmune disease;
KW brain tumour; brain injury.
XX

```



OS Homo sapiens.  
 XX MO200053774-A2.  
 XX 14-SEP-2000.  
 XX  
 XX 08-MAR-2000; 2000MO-US006237.  
 XX  
 XX 08-MAR-1999; 99US-00264585.  
 XX  
 XX (NEUR-) NEUROCRINE BIOSCIENCES INC.  
 XX  
 XX Kelnner GS, Clark M, Maki RA;  
 XX WPI; 2000-594326/56.  
 XX N-FSDB; AAA95821.  
 XX  
 XX Polynucleotide encoding novel members of a disintegrin, metalloproteinase  
 PT and thrombospondin domain protein family used to prevent and treat  
 PT Alzheimer's disease, cancer and autoimmune diseases.  
 XX  
 XX Claim 12; Fig 2; 129pp; English.  
 XX  
 XX The present sequence is human metalloproteinase ADAMTS-2. The ADAMTS  
 CC family of proteins is closely related to the ADAM (A Disintegrin and  
 CC Metalloproteinase Domain) family. Members of the ADAMTS family contain a  
 CC thrombospondin domain in addition to the disintegrin and  
 CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are  
 CC useful for the manufacture of medicaments for treating conditions  
 CC associated with neuroinflammation and/or neurodegeneration, such as  
 CC Alzheimer's disease, Parkinson's disease and stroke. They are also useful  
 CC for treating conditions associated with cell proliferation, cell  
 CC migration, inflammation and/or angiogenesis, such as cancer, arthritis  
 CC and autoimmune diseases. They can be used to treat patients afflicted  
 CC with an invasive tumour, a brain tumour or brain injury  
 CC  
 XX Sequence 680 AA;  
 SQ  
 Query Match 76.5%; Score 3713; DB 3; Length 680;  
 Best Local Similarity 99.1%; Pred. No. 9,4e-255;  
 Matches 674; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 XX  
 XX 211 RTKRFVSEARFVETLLVADASMAAFYGADLQNHILTLMSVARTYKPSIKNSINLWYK 270  
 DB 1 RTKRFVSEARFVETLLVADASMAAFYGADLQNHILTLMSVARTYKPSIKNSINLWYK 60  
 XX  
 XX 271 VLVDEKMGPEVSDNGGLTLNFCNMQRRPNQPSDRPHYDRAILLITRONPGCGGLC 330  
 DB 61 VLVDEKMGPEVSDNGGLTLNFCNMQRRPNQPSDRPHYDRAILLITRONPGCGGLC 120  
 XX  
 XX 331 DTLGVADIGTICDPNKSQVIEDEGLQAAHTLAHELGHTLSPHSDSKPCITRLFGPWGKI 390  
 DB 121 DTLGVADIGTICDPNKSQVIEDEGLQAAHTLAHELGHTLSPHSDSKPCITRLFGPWGKI 180  
 XX  
 XX 391 HTMAPEFTLNTQTLPMSPCSAMYLTELDDGSGGLLDAPAAALPLPTGLPGRAALYQLD 450  
 DB 181 HTMAPEFTLNTQTLPMSPCSAMYLTELDDGSGGLLDAPAAALPLPTGLPGRAALYQLD 240  
 XX  
 XX 451 QQCROIFGDFPHNCNTSADQVCAQLMCHTDGABPLCTKNGSLPMWDGTPCGGHLGSE 510  
 DB 241 QQCROIFGDFPHNCNTSADQVCAQLMCHTDGABPLCTKNGSLPMWDGTPCGGHLGSE 300  
 XX  
 XX 511 GSCLPHEVERKPYVDGGMAMPWGSGSRTCCGGVQVFSREKCKDEPONGRHYCLGRR 570  
 DB 301 GSCLPHEVERKPYVDGGMAMPWGSGSRTCCGGVQVFSREKCKDEPONGRHYCLGRR 360  
 XX  
 XX 571 AKYQSCHECECPDQSGFRQCEKXNANYTMDGNLLQVPPYAGVSPDRCKLFCRA 630  
 DB 361 AKYQSCHECECPDQSGFRQCEKXNANYTMDGNLLQVPPYAGVSPDRCKLFCRA 420  
 XX  
 XX 631 RGRSEKVFKAATVIGTLCGPETLACVRCQCVACGCHVDSRKLDPKCGCGKGNKSC 690  
 DB 421 RGRSEKVFKAATVIGTLCGPETLACVRCQCVACGCHVDSRKLDPKCGCGKGNKSC 480

XX 691 RKVSGSLPTNYGVNDIYITIPAGATNIDVKQRSHPGVQNDGNLYALATKADGOYLNGNLA 750  
 DB 481 RKVSGSLPTNYGVNDIYITIPAGATNIDVKQRSHPGVQNDGNLYALATKADGOYLNGNLA 540  
 XX  
 XX 751 ISAIQDILVKGTLTKSGSIATLERLOSRPPLPELTQLLTYPGVFPFKYTFEVP 810  
 DB 541 ISAIQDILVKGTLTKSGSIATLERLOSRPPLPELTQLLTYPGVFPFKYTFEVP 600  
 XX  
 XX 811 NDVDFSMQSKERATTNITQPLHAQWVLGDMSECSSTCCAGQRTVECRDPSGQASAT 870  
 DB 601 NDVDFSMQSKERATTNITQPLHAQWVLGDMSECSSTCCAGQRTVECRDPSGQASAT 660  
 XX  
 XX 871 CNKALKPEADAPCESQLCPL 890  
 DB 661 CNKALKPEADAPCESQLCPL 680  
 XX  
 XX RESULT 7  
 XX AAB21265  
 XX AAB21265 standard; protein; 896 AA.  
 XX  
 XX AAB21265;  
 XX  
 XX 23-FEB-2001 (first entry)  
 XX  
 XX Mouse metalloproteinase ADAMTS-1.  
 XX  
 XX Mouse; ADAMTS-1; metalloproteinase; ADAM;  
 KW a disintegrin and metalloproteinase domain; thrombospondin domain;  
 KW vaccine; neurotrophic; neuroprotective; antiparkinsonian; cerebroprotective;  
 KW cyostatic; antiarthritic; immunosuppressive; Alzheimer's disease;  
 KW Parkinson's disease; stroke; cancer; arthritis; autoimmune disease;  
 KW brain tumour; brain injury.  
 XX  
 XX Mus musculus.  
 OS  
 XX  
 XX MO200053774-A2.  
 XX  
 XX 14-SEP-2000.  
 XX  
 XX 08-MAR-2000; 2000MO-US006237.  
 XX  
 XX 08-MAR-1999; 99US-00264585.  
 XX  
 XX (NEUR-) NEUROCRINE BIOSCIENCES INC.  
 XX  
 XX Kelnner GS, Clark M, Maki RA;  
 XX WPI; 2000-594326/56.  
 XX  
 XX Polynucleotide encoding novel members of a disintegrin, metalloproteinase  
 PT and thrombospondin domain protein family used to prevent and treat  
 PT Alzheimer's disease, cancer and autoimmune diseases.  
 XX  
 XX Disclosure; Fig 17; 129pp; English.  
 XX  
 XX The present sequence is mouse metalloproteinase ADAMTS-1. The ADAMTS  
 CC family of proteins is closely related to the ADAM (A Disintegrin and  
 CC Metalloproteinase Domain) family. Members of the ADAMTS family contain a  
 CC thrombospondin domain in addition to the disintegrin and  
 CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are  
 CC useful for the manufacture of medicaments for treating conditions  
 CC associated with neuroinflammation and/or neurodegeneration, such as  
 CC Alzheimer's disease, Parkinson's disease and stroke. They are also useful  
 CC for treating conditions associated with cell proliferation, cell  
 CC migration, inflammation and/or angiogenesis, such as cancer, arthritis  
 CC and autoimmune diseases. They can be used to treat patients afflicted  
 CC with an invasive tumour, a brain tumour or brain injury  
 CC  
 XX Sequence 896 AA;  
 SQ  
 Query Match 48.3%; Score 2344; DB 3; Length 896;



```

QY 14 LILLILLPLARCAPAPPAAGQASBLVPT--RLPG-SAGELALHLSPKGFVRLA 70
DB 38 LILASITLILCARGAHGRFTE--EDELIVLPSELRAGHSTTRILDLDFGQDLHKIQ 95
QY 71 PDDSLAFEPFKIERLGGSGRATGGR-----GLRGCFSGTVNGEPESLAIVSICRG 123
DB 96 PDGSLAFGLFTLQIV---GRSPGSAQCLDPTDGLAHCFHGYNGVNGPSSAALSLCEV 152
QY 124 SSGFLDGEFTLQVGGAGSLAOPHRLQRMGPAPLPDGPWEVETGEG---ORGE 179
DB 153 RGAFLYLOEEFFIOPAP-----GVATERAAPVPEESSAPPOFHILRRR 198
QY 180 RGD-----HOEDSEESQEEAAGASEPPPLGA-----TSRTKRFVSEA 219
DB 199 RSGGAKCGVMDDELTPSDSPRESQNRNQMVRDPTPDAGKPSGSGIRKKRIVSAP 258
QY 220 RVEFTLLVADASMAAFYGADLONELITLMSVAARIYKPSIKXNINLMVYKVLIVEDK 279
DB 259 RVEFTLLVADQSMADFHSGSLKHYLLTFVVAARFYKPSIRNSISLVVYKVLIVIEBK 318
QY 280 GPEVNDNGGLTRNFCNNQRPNOPSDRPHRYDTAILLTRONCQOGLCDTLGVADIG 339
DB 319 GPEVSNALTLTRNFCNNQRPNOPSDRPHRYDTAILLTRONCQOGLCDTLGVADIG 377
QY 340 TICDPNKGCSVTEDEGLQAAHTLAELGHVLSMFDHDSKPTCLFPGMGHVMADPLFV 399
DB 378 TVCDPSRSCSVTEDEGLQAAHTLAELGHVLSMFDHDSKPTCLFPGMGHVMADPLFV 437
QY 400 INQTLPMSPCSAMYTELLDGGHGDCLDAPGALPLPTGPGMALYQDDQCRQIFPG 459
DB 438 LDHSGPMSPCSAMYTELLDGGHGDCLDAPGALPLPTGPGMALYQDDQCRQIFPG 495
QY 460 DRRHCENTSAQDVCAQMC-HTDGAEPICHTKNGSLPMADGTPCGFGLHSCBESCLPEE 518
DB 496 ESKHCPD--AASTCTTLMCTGSGGLVCCQTKH--FPMADGTPCGFGLHSCBESCLPEE 551
QY 519 VERPKPVVDGGAAPMGWGECSRTCGGVQPSHRECKDPEPONGRYCLGRAPKYOCSHT 578
DB 553 MGHFATPVHSGWGPMPGDDCSRTCGGVQPSHRECKDPEPONGRYCLGRAPKYOCSHT 611
QY 579 EBCPP-DGKSPFEOCCERKNAVNTDMGN--LLQWPKIAGVSPDRCLFPLANGRSE 635
DB 612 EBCPP-DGKSPFEOCCERKNAVNTDMGN--LLQWPKIAGVSPDRCLFPLANGRSE 670
QY 636 FVYFAKYIDGLCPETLALCVRGQCVKAGCDHVVSFRLDCKGVCGKNGSKRVSAG 695
DB 671 FVYFAKYIDGLCPETLALCVRGQCVKAGCDHVVSFRLDCKGVCGKNGSKRVSAG 730
QY 696 SILTPNYGVNDIVTPAGATNIDVKORSHPGVQNDGYALAKTDGQYILNGMLSAIE 755
DB 731 IYTSRTPGHVDIVTPAGATNIDVKORSHPGVQNDGYALAKTDGQYILNGMLSAIE 790
QY 756 QDILVKGITLKXSGSIATLERLQSEFRLPEBELTYQLTVGEVFPKAKYTFVPVNDV 815
DB 791 QDILVKGITLKXSGSIATLERLQSEFRLPEBELTYQLTVGEVFPKAKYTFVPVNDV 843
QY 816 SMOSSKERATNIIOPPLHAQVTLGMEGCSITGAGWQRTTYEGRPSQAATCNKL 875
DB 844 ---KKKTESFVAITFT--SEMVIEMGECSTKSGQRRVAVQCRDINGHPABEQAKEY 897
QY 876 KPEDAKCESQCLP 889
DB 898 KPASTRPCADLP 911

```

```

RESULT 9
AA49501
ID AA49501 standard; protein; 950 AA.
XX AA49501;
XX AA49501;
DT 10-JAN-2000 (first entry)
XX

```

```

DE Human METH1 protein.
XX Human; METH1; METH2; anti-angiogenic; metalloproteinase thrombospondin;
XX cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
XX angiogenesis inhibitor; abnormal wound healing; inflammation;
XX rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
XX diabetic retinopathy; macula degeneration; hemangioma; detection;
XX arterial-venous malformation; immune deficiency.
XX Homo sapiens.
XX WO937660-A1.
XX 29-JUL-1999.
XX 22-JAN-1999; 99MO-US001313.
XX 23-JAN-1998; 98US-0072288P.
XX 28-AUG-1998; 98US-0098539P.
XX (IRUE/) IRUELA-ARISPE L.
XX (HAST/) HASTINGS G A.
XX (RUBE/) RUBEN S M.
XX IrueLA-Arispe L, Hastings GA, Ruben SM,
XX WPI, 1999-590684/50.
XX N-PEDB; AA232000.
XX New isolated metalloproteinase thrombospondin polypeptides, useful for
XX treating hyperproliferative disorders, cancers or autoimmune disorders.
XX Claim 10; Fig 1; 457pp; English.
XX AA232000 and AA232001 encode, and AA49501 and AA49502 represent, human
XX metalloproteinase thrombospondin (METH) proteins METH1 and METH2
XX respectively. METH1 and METH2 have been found to be potent inhibitors of
XX angiogenesis both in vitro and in vivo. They can be used for treating
XX cancer and other disorders related to angiogenesis including abnormal
XX wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial
XX bleeding disorders, diabetic retinopathy, some forms of macula
XX degeneration, haemangiomas, and arterial-venous malformations. They may
XX be useful in treating deficiencies or disorders of the immune system, by
XX activating or inhibiting the proliferation, differentiation, or
XX mobilization (chemotaxis) of immune cells. The etiology of these immune
XX deficiencies or disorders may be genetic, somatic, such as cancer or some
XX autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or
XX infectious. They can also be used to treat inflammatory conditions, both
XX chronic and acute conditions. The products can also be used for detection
XX and diagnosis. AA232002 to AA232080, and AA49503 to AA49511 represent
XX sequences given in the exemplification of the present invention
XX
XX Sequence 950 AA;
XX
XX Query Match 48.0%; Score 2328.5; DB 2; Length 950;
XX Best Local Similarity 49.7%; Pred. No. 2,7e-156;
XX Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;
XX
XX 11 LPTLLILLPLARCAPAPPAAGQASBLVPT--TRLPGSAGELALHLSPKGFVRLA 68
XX 16 VPTLLILAAALAVS--DALGRPSE--EDELIVLPSELRAGHSTTRILDLDFGQDLHKIQ 71
XX 69 LAPDSEFLAFEPFKIERLGGSGRATGGR-----ERGLGCFSGTVNGEPESLAIVSICRG 123
XX 72 LAPDSEFLAFEPFKIERLGGSGRATGGR-----ERGLGCFSGTVNGEPESLAIVSICRG 128
XX 124 SSGFLDGEFTLQVGGAGSLAOPHRLQRMGPAPLPDGPWEVETGEG---ORGE 172
XX 129 RGAFLYLOEEFFIOPAP-----GVATERAAPVPEESSAPPOFHILRRR 183
XX 173 GEGORERGDHSEESQEEAAGASEPPPLGA-----TSRTKRFVSEA 222
XX 184 VDDEPRTPGAEFTEDDEDEGTEGDEBPGWSPDPAQGVGQPTGTGTSIKKGFVSSHRV 243

```

QY 223 ETLVADASMAAFYADLQNHILTLMSVAARIYKPSIKNSINIMVVKYLVEDEKMGPE 282  
 DB 244 ETLVADQSMAPFSGSLKHYLLTLFVAARLYGSPISIRNSVSLVVKYLIVHDEKMGPE 303  
 QY 283 VSDNGGLTLRNFQCMQORRPNQSDRPHENYDTALLTRQNFQCGQEGCLDTLGVADIGTIC 342  
 DB 304 VTSNAALTLRNFQCMQORRPNQSDRPHENYDTALLTRQNFQCGQEGCLDTLGVADIGTIC 362  
 QY 343 DPNKSCSVIEDBGLQAAATLAHELGHVLSMPHDSKPCRTLFGPMGKHVYADLVFVHLNQ 402  
 DB 363 DPNKSCSVIEDBGLQAAATLAHELGHVLSMPHDSKPCRTLFGPMGKHVYADLVFVHLNQ 422  
 QY 403 TLPMSPGSAWYLTLLDGGHDDCLLDAPGALPLPTGLPGPMALYUDDQCCROIFPDPDR 462  
 DB 423 SPMSPGSAWYLTLLDGGHDDCLLDAPGALPLPTGLPGPMALYUDDQCCROIFPDPDR 480  
 QY 463 HCPNTSAQDVCAQLWC-HTDGAEPFLCHTKNGSLPMWADGTFCGPGHLCSEGSCLPEEVEVER 521  
 DB 481 HCPD--AASCTSLMCTGTSGVLVCOYTKH--FPWADGTSCGEGKWCINGKCVNKTDRKH 536  
 QY 522 PKPVVDGGMAPMGWGCSTRTCGGVOFPHRECKDEPBGNGRYCGRAXKYOSCHTEEC 581  
 DB 537 FDTFPHSGMGMWGCSTRTCGGVOFPHRECKDEPBGNGRYCGRAXKYOSCHTEEC 596  
 QY 582 PP-DGKSFREOQCEKYNAYNTDM-DGNLQWPKYAGVSPRDRCKLFCRARGSBPKYF 639  
 DB 597 PDNNGKTFREOQCEKYNAYNTDM-DGNLQWPKYAGVSPRDRCKLFCRARGSBPKYF 656  
 QY 640 EAKYIDGLTLCGPETLAIQVGGQCVKAGCDHVNSPRRLDKQVCGGKNSCRVSGSLTP 699  
 DB 657 QPKYVDGTPGSPDSTVCGQCVKAGCDHVNSPRRLDKQVCGGKNSCRVSGSLTP 716  
 QY 700 TNYGNDIVITI PAGATINDVQRSHPVQNDGNVLAUKLADGQYLLNGMLAISABODIL 759  
 DB 717 AKPGYHDIITITPGATINIEVKQRNORSGRNNGSFLAIKADGTYLLNGDYTLSTLBDIM 776  
 QY 760 VKGTLKXSGSLATLERQSPRLPEPLTVOLLVPEBVPKPKYTFPFVNDVDSMGS 819  
 DB 777 YKGVNLRTSGSALERTSRSPFLKEPLTIDVLTIV-GNALRPKIKTYFV----- 825  
 QY 820 SKERATITIIOPLLAQWLVLDWMSGCSSTCGAGQORRTVECRDPSGQASATCNKALKPBD 879  
 DB 826 KKKKSPNAL-PTFSA-WVIEWEGCSKSCELGQRLVCECRDINGPASECAKAYKPAS 883  
 QY 880 AKPCESQICP 889  
 DB 884 TRPCADHPCP 893

RESULT 10  
 AAB73549  
 ID AAB73549 standard; protein; 950 AA.

AC AAB73549;  
 DT 07-AUG-2001 (first entry)  
 DE Human ADAM-type metalloprotease MDTs4, SEQ ID NO:4.  
 KW Human; MDTs4; ADAM-type metalloprotease; drug screening;  
 XX A Disintegrin And Metalloprotease; cancer; arthritis.  
 OS Homo sapiens.  
 PN JP2001017183-A.  
 PD 23-JAN-2001.  
 PF 09-JUL-1999; 99JP-00196584.  
 PR 09-JUL-1999; 99JP-00196584.

PA (YAMA) YAMANOCHI PHARM CO LTD.  
 DR WPI; 2001-275950/29.  
 DR N-PSDB; AAH20224.  
 XX A new metal protease and its preparation for use as an anti-cancer and  
 PT anti-arthritis therapeutic.  
 PS Claim 1; Page 12-14; 22pp; Japanese.  
 CC The invention relates to the novel human ADAM (A Disintegrin And  
 CC Metalloprotease)-type metalloproteases MDTs4 (AAB73549) and MDTs5  
 CC (AAB73550). The metalloproteases can be used for the treatment of cancers  
 CC and arthritis. The invention also relates to the genes encoding MDTs4 and  
 CC MDTs5, vectors and host cells containing the MDTs4 or MDTs5 genes, the  
 CC recombinant production of MDTs4 and MDTs5, and antibody specific for  
 CC MDTs4 or MDTs5, and methods of screening for compounds which modulate the  
 CC activity of MDTs4 and/or MDTs5. The present sequence represents human  
 CC MDTs4  
 SQ Sequence 950 AA;  
 Query Match 48.0%; Score 2328.5; DB 4; Length 950;  
 Best local similarity 49.7%; Pred. No. 2.7e-156;  
 Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;  
 11 LPLULLLULLLPLARGAPAPPAQASLVP--TRLSAGSLALHLISAFKGFVLR 68  
 16 VPTLLLAALLAVS-DALGRPSE--EDELVYPELEBARPGH-GTYRLHLHFDQDLDE 71  
 69 LAPDPSFLAPEFKERLGGSGRANG-----ERGLAGCFPGSGTVNGEPSSIAVSLCRGL 123  
 72 LRPDSFLAPGFTLQNV--GRKSGETPLPETHLHCHTSGVINGDPSAALLSLCEGV 126  
 124 SGSFLDGEETFIOPQAGGSLAOPHLQRMGPAGARPLP-----RGPWEV---ET 172  
 129 RGAFTYLLGEAVFIQPLDPAAS-----ERLATPAGKEKPPAPLQFHLRRRQGVGGTGV 183  
 173 GEGROGERGHQDESE 222  
 184 VDDEFRPTGAETDEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDE 243  
 223 ETLVADASMAAFYADLQNHILTLMSVAARIYKPSIKNSINIMVVKYLVEDEKMGPE 282  
 DB 244 ETLVADQSMAPFSGSLKHYLLTLFVAARLYGSPISIRNSVSLVVKYLIVHDEKMGPE 303  
 QY 283 VSDNGGLTLRNFQCMQORRPNQSDRPHENYDTALLTRQNFQCGQEGCLDTLGVADIGTIC 342  
 DB 304 VTSNAALTLRNFQCMQORRPNQSDRPHENYDTALLTRQNFQCGQEGCLDTLGVADIGTIC 362  
 QY 343 DPNKSCSVIEDBGLQAAATLAHELGHVLSMPHDSKPCRTLFGPMGKHVYADLVFVHLNQ 402  
 DB 363 DPNKSCSVIEDBGLQAAATLAHELGHVLSMPHDSKPCRTLFGPMGKHVYADLVFVHLNQ 422  
 QY 403 TLPMSPGSAWYLTLLDGGHDDCLLDAPGALPLPTGLPGPMALYUDDQCCROIFPDPDR 462  
 DB 423 SPMSPGSAWYLTLLDGGHDDCLLDAPGALPLPTGLPGPMALYUDDQCCROIFPDPDR 480  
 QY 463 HCPNTSAQDVCAQLWC-HTDGAEPFLCHTKNGSLPMWADGTFCGPGHLCSEGSCLPEEVEVER 521  
 DB 481 HCPD--AASCTSLMCTGTSGVLVCOYTKH--FPWADGTSCGEGKWCINGKCVNKTDRKH 536  
 QY 522 PKPVVDGGMAPMGWGCSTRTCGGVOFPHRECKDEPBGNGRYCGRAXKYOSCHTEEC 581  
 DB 537 FDTFPHSGMGMWGCSTRTCGGVOFPHRECKDEPBGNGRYCGRAXKYOSCHTEEC 596  
 QY 582 PP-DGKSFREOQCEKYNAYNTDM-DGNLQWPKYAGVSPRDRCKLFCRARGSBPKYF 639  
 DB 597 PDNNGKTFREOQCEKYNAYNTDM-DGNLQWPKYAGVSPRDRCKLFCRARGSBPKYF 656  
 QY 640 EAKYIDGLTLCGPETLAIQVGGQCVKAGCDHVNSPRRLDKQVCGGKNSCRVSGSLTP 699  
 DB 657 QPKYVDGTPGSPDSTVCGQCVKAGCDHVNSPRRLDKQVCGGKNSCRVSGSLTP 716

QY 700 TNYGNDIVTTTATGATNIDYKORSHPVQNDGNATLAKTADQTLNGNLAIKIEDIL 759  
 DB 717 AKEGHDTITPTGATNIEVKQNRGSRNNGSFLAIAAGTYILNGDYLTSTLEDDIM 776  
 QY 760 VKGTILKSGSITATLERTLQSFRLPEPLTVOLLTVPGEVFPFKYKTYFEVNDVDFSMQS 819  
 DB 777 YKGVILRYSGSSAALERIRSPSPLEPILITVLT-QUALPKIKTYFV----- 825  
 QY 820 SKERATTNIOPLIAQWTLGPMWSSCSTCCAGMGRRTVERDRDSGQASATCKALKPED 879  
 DB 826 KKKKSEFNAL-PTFSA-WVIEBEGCSKSCELGWRRLVEGRDINGQDPAECAKEVPAS 883  
 QY 880 AKCESQLCP 889  
 DB 884 TRPCADHPCP 893

RESULT 11  
 AAB50002  
 ID AAB50002 standard; protein; 950 AA.  
 AC AAB50002;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE Human METH1.  
 XX  
 KW Human; METH1; metalloprotease; thrombospondin; angiogenesis inhibition;  
 KW cancer therapy; benign tumour; ocular angiogenic disease;  
 KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;  
 KW vasculodermis; granulation; hypertrophic scar; nonunion fracture;  
 KW scleroderma; trachoma; vascular adhesions; myocardial angiogenesis;  
 KW coronary collateral; cerebral collateral; arteriovenous malformation;  
 KW ischaemic limb angiogenesis; Osler-Webber syndrome;  
 KW plaque neovascularisation; telangiectasia; haemophilic joint;  
 KW Crohn's disease; fibromuscular dysplasia; wound granulation;  
 KW Crohn's disease; atherosclerosis; birth control.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200071577-A1.  
 XX  
 PD 30-NOV-2000.  
 XX  
 PF 25-MAY-2000; 2000WO-US014462.  
 XX  
 PR 25-MAY-1999; 99US-00318208.  
 PR 20-JUL-1999; 99US-0144882P.  
 PR 10-AUG-1999; 99US-0147823P.  
 PR 13-AUG-1999; 99US-00373658.  
 PR 22-DEC-1999; 99US-0171503P.  
 PR 22-FEB-2000; 2000US-0183792P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (SMK-) SMITHKLINE BEECHAM CORP.  
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
 PA (IRUE/) IRUELA-ARISPE L.  
 PA (HAST/) HASTINGS G A.  
 PA (RUBE/) RUBEN S M.  
 PA (JONK/) JONAK Z L.  
 PA (TRUL/) TRULLI S H.  
 PA (FORN/) FORNWALD J A.  
 PA (TERR/) TERRETT J A.  
 XX  
 PI IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH,  
 PI Fornwald JA, Terrett JA;  
 XX  
 DR WPI, 2001-025136/03.  
 DR N-PDB; AAC90057.  
 XX  
 PT METH1 and METH2 polynucleotides and encoded polypeptides, used to inhibit  
 PT angiogenesis in the treatment of disorders such as cancer, rheumatoid

PT arthritic and psoriasis.  
 XX  
 XX Claim 15, Fig 1, 768pp; English.  
 XS  
 CC The present sequence is human METH1 (ME for metalloprotease and TH for  
 CC thrombospondin). METH1 can be used for inhibiting angiogenesis in an  
 CC individual, and for treating cancer, benign tumours, an ocular angiogenic  
 CC disease, rheumatoid arthritis, psoriasis, delayed wound healing,  
 CC endometriosis, vasculogenesis, granulations, hypertrophic scars, nonunion  
 CC fractures, scleroderma, trachoma, vascular adhesions, myocardial  
 CC angiogenesis, coronary collateral, cerebral collateral, arteriovenous  
 CC malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiodioma, plaque  
 CC fibromuscular dysplasia, wound granulation, Crohn's disease or  
 CC atherosclerosis. METH1 can also be used in birth control. METH1 can also  
 CC be used in diagnostic methods for the prognosis of cancer  
 CC  
 SO Sequence 950 AA;  
 Query Match 48.0%; Score 2328.5; DB 4; Length 950;  
 Best Local Similarity 49.7%; Pred. No. 2,7e-156;  
 Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;  
 QY 11 LPFLILLILLILPLARAGAPAPAGQASELVVP--TLPGSAGELALHLSAFKGVLR 68  
 DB 16 VPTILLILAAALILAVS-DALGRPS--EDELIVPELERARGH-GTTRLRILHARDQDLDE 71  
 QY 69 LAPDSFLAPFEKTERIGSGGRATG3-----ERLRCGFPSTGVNGBPSLAVALICRL 123  
 DB 72 LRPSFFLAPFTLQNV---GRKSGSTPLPEETDLAHCFYSTGVNBPSSAALSLCEGV 128  
 QY 124 SSGFLDDEERTTIOPOGAGSLAQPHRLQWGPAGAPLP-----RGPEV---ET 172  
 DB 129 RGATFLGEATFIQPLPAAS-----EKLATRAQEKRPALQHLRRNQGVGATCGV 183  
 QY 173 GEGQGERGHQEDSEESCEERAGASEPP-----PIGATLS-RKRVASARFY 222  
 DB 184 VDDEPRPFGKAEETDEBEGTEDEGPQWSPQDPAQGVQFPGTGSIRKRVSSHRY 243  
 QY 223 ETLVADASMAFYGADLQNLITLMSVAARIYKHPISIKSINLMVKVLIVDEKMGPE 282  
 DB 244 ETMLVADQSMAPFHSGSLKHLITLFSVAAKLTKHPSIRNSVSLVYKILVHDEKGP 303  
 QY 283 VSDNGGLTLRNFCWQRFRNQPSPRPHRYDTAILLTRQNFCCQBLCDTLGVADIGTIC 342  
 DB 304 VTSNAALTLRNFCWQKQHPSPDRDAEHYDTAILLTRQDLCSQ-TCDTLGVADIGTIC 362  
 QY 343 DENKCSYIEDEGLQAAHTLAHELGHVLSMHPDCKPCTRLFGPMGKHVMARLPVHLNQ 402  
 DB 363 DPSRCSYIEDDGLQAAFTTHHELGHVFMNPHDQKQASLVNGQDSHMAAMLNLIDH 422  
 QY 403 TLPMSPGSAWYLTLELDGSHDCLIDPAAALPLPTGLPGRMALYQIDOCROTLFGPDFR 462  
 DB 423 SOPMSPCSAWYLTLELDGSHDCLIDPAAALPLPTGLPGRMALYQIDOCROTLFGPDFR 480  
 QY 463 HCPNTSAQDVQALWC-HTDGAZPLCHTKNGSLPMDADTCGGRHLCSGSCULPEEVER 521  
 DB 481 HCPD--AASTGSLTMCGTSGGLVWCOTKH--FPMADGTCGGRKWCINCKVAKTRKH 536  
 QY 522 PKPVVDGMAWPMWGBESRTCCGGGVOPSHRECDPCKNGSVYCGRAKYOSCTHEEC 581  
 DB 537 FDTFPHGSMGMPWGDSTRTCGGGVYTRBCDNPVPAKGGKTCDEKRVRYSCNLEDC 596  
 QY 582 PP-DGKSPREOCCERKXNAUNYTDN-DGNLLQWPKYAGVSPDRCKLFCRAGRSEFKYF 639  
 DB 597 PDNNGKTRRESCCAHNEFSKASFGSPAYEWMPKXGAVSPKORCKLICAKIGYFFVL 656  
 QY 640 EAKYVDGFLCGPBTALCVRGQCVKAGCDVYVSPRLTDCGCVGGKSGCRKVGSLTP 699  
 DB 657 QPKVVDGTPCSPDSTSVQVQCVKAGCDRIIDSKKGFDCGCVGGKSGCRKISGYS 716  
 QY 700 TNYGNDIVTTTATGATNIDYKORSHPVQNDGNATLAKTADQTLNGNLAIKIEDIL 759

Db 717 AKPGYHDIITITPTGATNIEVKQNRGSRNNGSFLAIKADGTYIINGDYITLSTLEODIM 776  
 QY 760 VKGTLIKYSGSIATLERLOSFRPLPEPLTVQLLTVGEVFPKVKYTFVNDVDFSMOS 819  
 Db 777 YKGVVLRYSGSALERIRSPSPLEKEPLTIQVLTV-GNALRPKIKYTFV----- 825  
 QY 820 SKERATNTIITQPLHQAQVLCGDMSECSSTCGAGMQRRTVCRPSCQASATCKALKPED 879  
 Db 826 KKKKESFNAL-PTPSA-WVLEWGECSKCELGMRRLVBCRDINGQAPASECKAEVKPAS 883  
 QY 880 AKPCESQLCP 889  
 Db 884 TRPCADHPCP 893  
 RESULT 12  
 AAW80285 standard; protein; 967 AA.  
 ID AAW80285 standard; protein; 967 AA.  
 AC AAW80285;  
 XX 19-JAN-1999 (first entry)  
 DT 19-JAN-1999 (first entry)  
 DE Human integrin ligand polypeptide ITGL-TSP.  
 KW ITGL-TSP; integrin ligand; angiogenic disease; cancer; atherosclerosis;  
 chronic inflammatory disorder; rheumatoid arthritis; tissue remodeling;  
 macular degeneration; diabetic retinopathy; Alzheimer's disease; human;  
 restenosis.  
 KM Homo sapiens.  
 OS Homo sapiens.  
 PN EP874050-A2.  
 PD 28-OCT-1998.  
 PE 27-JAN-1998; 98EP-00300575.  
 PR 24-APR-1997; 97US-00845496.  
 PA (SMIK) SMITHKLINE BEECHAM CORP.  
 PA (HUKA-) HUMAN GENOME SCI INC.  
 PI Jonak ZL, Trullis SH, Fromwald JA, Hastings GA, Terrett JA.  
 DR WPI, 1998-544643/47.  
 DR N-PEDB; AAV66508.  
 PT DNA encoding integrin ligand polypeptide ITGL-TSP - used to treat  
 angiogenic diseases, restenosis, Alzheimer's disease and in tissue  
 remodeling.  
 PS Claim 11, Page 6-9; 24pp; English.  
 CC This represents a human integrin ligand polypeptide ITGL-TSP. ITGL-TSP  
 polypeptides can be used in the treatment of angiogenic diseases such as  
 cancer, cancer metastasis, chronic inflammatory disorders, rheumatoid  
 arthritis, atherosclerosis, macular degeneration or diabetic retinopathy,  
 restenosis, Alzheimer's disease and tissue remodeling. They can be used  
 to treat a subject in need of enhanced activity or expression of the ITGL-  
 TSP polypeptide  
 CC  
 CC  
 SQ Sequence 967 AA;

Query Match 48.0%; Score 2328.5; DB 2; Length 967;  
 Best Local Similarity 49.7%; Pred. No. 2.8e-156;  
 Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

QY 11 LPFLILLILLPLARGPAPAPAGQASELYVP--TRLGSGAGELALHLSAFQKGFYLR 68  
 Db 33 VPTLLLAALAAVS-DALGRPSE--EDDELVPELERAPGH-GTTRLRLHAFDQDLDE 88

QY 69 LAPDSELADEFKIERLGGGRATGG-----ERGLNGCFSGITVNGEPESLAASLCRGL 123  
 Db 89 LRPFSSFLAAGFTLQNV---GRKSGSEPLPEFDLHNCYSGITVNGDPSSAALSLICEYV 145  
 QY 124 SGEFLDGESEFTIOPQAGGSLAQPHRLQRMWGPAGARPLP-----RGSEMEV-----ET 172  
 Db 146 RGAFTYLIGEVYFQPLPAAS-----ERLATAPAGEPRPPLQPHLLRRRQDVGCTGV 200  
 QY 173 GSGQRQKRGHQDSEBESOEBAEGASEPP-----PLGATS-RTKRFVSDARFV 222  
 Db 201 VDDEPRPTGAKETDEDEGEDEBQPOWSPPDPALQVGPFTGTSIRKKRFVSHRTV 260  
 QY 223 ETLIVADASMAAFYADLQNHILTLMSVAARLYKPSIKNSINLWVKYLIYEDSKMGP 282  
 Db 261 ETMLVADQSMAEFHSGGLKHYLLTLFSVAARLYKPSINSVSLVYVKILVTHDQKGE 320  
 QY 283 VSDNGSLTLRNFQNRFRNPQPSDRHPEHYDTALILTRONFCQGEGLCDTLGVADIGTC 342  
 Db 321 VTSNAALTLENFCNQKQHNPPSDRDAREHYDTALIFTRQDLGSGQ-TCDTLGMADVGTVC 379  
 QY 343 DPNKSCSVIEDEGLQAAHTLAHELGHVLSMPHDDSKPCTRIRFGPMKHHVMAPLFVHLNQ 402  
 Db 380 DSRKSCSVIEDGLQAAFTTAHELGHVFNPHDDAQCSLNGVANDSMMSMLSLNDH 439  
 QY 403 TLPWSPCSAMYLTLLDGHGHDCLDAPALPLPRTLPGSMALYQLDQCGQIFQDPFR 462  
 Db 440 SQPSPCSAVYITSLFDNGHGECLMDKPNPILPQDLPG--TSYDANRQCQFTGEDESK 497  
 QY 463 HCPRTSAQDVCAQLWC-HTDGAEPILCHTNKGLPAMDGPCCGPHLCSGSGCLPEBEVR 521  
 Db 498 HCPD--AASICSFLMCTGTSGLVLCQTKH--FPMAHGISCGBKMCITGKVNKTRGH 553  
 QY 522 PKPVVDGMAFWGPMGECSTFCGGVQVFSHRECKDPEPONGRYCLGRPAKYQSCHTEEC 581  
 Db 554 FDFEFHSGMWGMGMPGDCSFTCGGVQYTMREBDNVPNNGGKYCBGKVRYSNLEBDC 613  
 QY 582 PP-DGKSFREQCEKKNAYVYTDV--DGNLLQWPKYAGVSPDRCKLFCRARGSEPKYF 639  
 Db 614 PDNNGKTFREQCEANHEFSKASFGSGPRAVENIPKTAGVSPDRCKLQCAKGTGTFVL 673  
 QY 640 EAYVIDTGLCPETLAIYRGQCVKAGCDHVVDSPKLDKCYVCGGKNSCKRVGSLTP 699  
 Db 674 QPKVVDGTPCSPDSTVCVQGCVKACCDRIIDSKKPKFCVCGGNGTCKKIGSAYTS 733  
 QY 700 TNYGNDIVYIPAGATNIDVQKSHRPGVQNDGNYALKTAPDQVYLINGLALSALEODIL 759  
 Db 734 AKPGYHDIITITPTGATNIEVKQNRGSRNNGSFLAIKADGTYIINGDYITLSTLEODIM 793  
 QY 760 VKGTLIKYSGSIATLERLOSFRPLPEPLTVQLLTVGEVFPKVKYTFVNDVDFSMOS 819  
 Db 794 YKGVVLRYSGSALERIRSPSPLEKEPLTIQVLTV-GNALRPKIKYTFV----- 842  
 QY 820 SKERATNTIITQPLHQAQVLCGDMSECSSTCGAGMQRRTVCRPSCQASATCKALKPED 879  
 Db 843 KKKKESFNAL-PTPSA-WVLEWGECSKCELGMRRLVBCRDINGQAPASECKAEVKPAS 900  
 QY 880 AKPCESQLCP 889  
 Db 901 TRPCADHPCP 910

RESULT 13  
 AAB50011 standard; protein; 968 AA.  
 ID AAB50011 standard; protein; 968 AA.  
 AC AAB50011;  
 XX 19-MAR-2001 (first entry)  
 DT 19-MAR-2001 (first entry)  
 DE Protein; SEQ ID 125.  
 KM Human; MEN1L, metalloprotease; thrombospondin; angiogenesis inhibition;  
 cancer therapy; benign tumour; ocular angiogenic disease;



KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;  
 KW vasculogenesis; granuloma; hypertrophic scar; nonunion fracture;  
 KW scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;  
 KW coronary collateral; cerebral collateral; arteriovenous malformation;  
 KW ischaemic limb angiogenesis; Osler-Weber syndrome;  
 KW plaque neovascularisation; telangiectasia; haemophilic joint;  
 KW angiodysplasia; fibromuscular dysplasia; wound granulation;  
 KW Cronin's disease; atherosclerosis; birth control.  
 OS Homo sapiens.  
 XX WO200071577-A1.  
 XX 30-NOV-2000.  
 XX 25-MAY-2000; 2000WO-US014462.  
 XX 25-MAY-1999; 99US-00318208.  
 XX 20-JUL-1999; 99US-0144882P.  
 XX 10-AUG-1999; 99US-0147823P.  
 XX 13-AUG-1999; 99US-00373658.  
 XX 22-DEC-1999; 99US-0171503P.  
 XX 22-FEB-2000; 2000US-0183792P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX (SMK) SMITHKLINE BEECHAM CORP.  
 XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
 XX (IRUE/) IRUELA-ARISPE L.  
 XX (HAST/) HASTINGS G A.  
 XX (RUBEN/) RUBEN S M.  
 XX (JONAK/) JONAK Z L.  
 XX (TRULI/) TRULI S H.  
 XX (FORN/) FORN W D.  
 XX (TERRETT/) TERRETT J A.  
 XX IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;  
 XX Formwald WJ, Terrett JA.  
 XX WPI, 2001-025136/03.  
 XX METH1 and METH2 polynucleotides and encoded polypeptides, used to inhibit  
 XX angiogenesis in the treatment of disorders such as cancer, rheumatoid  
 XX arthritis and psoriasis.  
 XX Claim 15; Page 759-763; 768pp; English.  
 CC The present invention relates to human METH1 and METH2 (ME for  
 CC metalloproteinase and TH for thrombospondin; see AAB5002 and AAB5003).  
 CC METH can be used for inhibiting angiogenesis in an individual, and for  
 CC treating cancer, benign tumours, an ocular angiogenic disease, rheumatoid  
 CC arthritis, psoriasis, delayed wound healing, endometriosis,  
 CC vasculogenesis, granuloma, hypertrophic scars, nonunion fractures,  
 CC scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,  
 CC coronary collaterals, cerebral collaterals, arteriovenous malformations,  
 CC ischaemic limb angiogenesis, Osler-Weber syndrome, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiodysplasia,  
 CC fibromuscular dysplasia, wound granulation, Cronin's disease or  
 CC atherosclerosis. METH can also be used in birth control. METH can also be  
 CC used in diagnostic methods for the prognosis of cancer. The present  
 CC sequence is a protein isolated in the present invention  
 XX Sequence 968 AA;  
 SQ  
 Query Match 48.0%; Score 2328.5; DB 4; Length 968;  
 Best Local Similarity 49.7%; Pred. No. 2.8e-156;  
 Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;  
 QY 11 LPELLILLILLIPAGAPAPAGAGAGSEVVP--TLPSAGELALHLSAFGKGFVLR 68  
 DB 34 VPTLLILLALLAVS-DALGRPS--EDBELVPELEBAPPH-GTTLLRLHAFQGDLE 89  
 QY 69 LAPDSFLAPERIERLGGSGRATG-----ERGLRGCFSGTNGEPESLAAYSLCRGL 123

DB 90 LRPDSFLAPGFTLQNV--GRKSGSETPLPETDLACHFSGTVNGDPSSAALLSLCEGV 146  
 QY 124 SGTFLDGEERTIOPQAGAGSLAPPHLQGWGAGAPLP-----RGPEWEV-----ET 172  
 DB 147 RGAIVYLGENAVFIQPLPAS-----ERLAPAAAGEPPAPLQHLRRRRQGVGTGCV 201  
 QY 173 GEGORGERGDHQBESBESQEBEAGASEPP-----PPGANS-RTKPFVSEARV 222  
 DB 202 VDDEPRPTGAEETDEDETEGEGEDGQWSPDPALQGVQPTGTSIRKKPFVSSHRV 261  
 QY 223 ETLVADASMAAFYADLQNHILTLMSVARIYKFSIKNSIMLVYKVIYEDKMGPE 282  
 DB 262 ETMLVADQSMAPFGSGSLKHILTLFVARLKHFSINSVSLVYKVLIVHDEKXGE 321  
 QY 283 VSDNGGLTLRNFPCWQRRFNPQSDRPHRYDTNLLITRONFGQBELCOTLGVADIGTIC 342  
 DB 322 VTSNAALTLRNFQWQKQHPPSDRDAEHYDTALFTRQDLQSQ--TCDTLQVADGVIC 380  
 QY 343 DENKSGVIEDEGLQAAHTLAHELGHVLSMHDSPKCTELFGPMKHHVMAPLFVHLNQ 402  
 DB 381 DSRSCSVIEDDGLQAAFTTAHELGHVFMPHDARQASLNGVNDSHMMSMLNDH 440  
 QY 403 TLPWSPGSAIYLTBLDGGHGDCLDAPGAALPLFTGLFGKRALYQLDQCCROIPEPDR 462  
 DB 441 SQPWSFSAVMTISFLDNGHGECLMDKPPNFIQLPDLRQ--TSYDANQCCPTFEDSK 498  
 QY 463 HCPNTSADVQACLMC-HTDGAEPCLCTKNGSLPMADGTPCGPHLCSGSLPEBEVER 521  
 DB 499 HCPD--AASTCSITLWCSTGSGVLVQTKH--FPNADGTSCEGKMKCINGKCNKTRDGH 554  
 QY 522 PKPVVDGMAWPGWGCSTCCGGVQFSTRCKDPEPONGARYCLGRAPKQSCITEC 581  
 DB 555 FDTPFHSGWGMGPGSCSPTCCGGVQYTRBEDNVPKNGKRYCKGKRYRSCNLEDC 614  
 QY 582 PP-DKSFREQCEKRYAAYVYTDN-DGNLLQWPKTAGVSPEDRCCLFRRAGRSEFKVF 639  
 DB 615 PDNNKTFREBQCEAHNEFKAQSPGSPAYEMVLPKAGVSPDRCLTLQAGIGVFVL 674  
 QY 640 EAKYIDGTLGPELTALCVRGCCYKAGCGHNVDSPKLDCGVCCKGKNSCRVSGSLP 699  
 DB 675 QPKYVDGTPSCSDTSCVQGCYKAGCDRIIDSKKPKCGVCGGNGSTCKKISSVTS 734  
 QY 700 TNYGNDIVTIPAGATNIDVKQSHSPGVNDGNVLAKTADGGYLLNGMLASALBODIL 759  
 DB 735 AKPEVHITITPTCATNIEVKQNRNGSRNNGSFLAKADGVYLLNGYTLSTLEODIM 794  
 QY 760 VKGTLIKYSGSIATLELQSPRPPLPEPLTYQLITVGEVPPPKVKTTFVENDVDSMG 819  
 DB 795 YKGVILRYSGSALERIRFSPLEKPLTIOVLV--GNALRPKIKTYFV----- 843  
 QY 820 SKERATNIIPLHQAQVLDWSECSSTGACQMRRTVECRDPSCQASATCNKALKPD 879  
 DB 844 KKKESFPAI-PTPSA-VNIEWGECKSCCELQMRVLVECRDINGQPASECAYKVPAS 901  
 QY 880 AKPGESQLCP 889  
 DB 902 TRPCADHPCP 911  
 RESULT 14  
 ID AAY04142  
 AA AAY04142 standard; protein; 967 AA.  
 XX AC AAY04142;  
 XX DT 15-JUN-1999 (first entry)  
 XX DE Human TANGO-71 protein.  
 XX KM Human; TANGO-71; TANGO-73; TANGO-74; TANGO-76; TANGO-83; diagnosis;  
 XX KM detection.  
 OS Homo sapiens.





CC 269-456. In *C. elegans* hermaphrodites, GON-1 is required for migration of  
CC two distal tip cells to produce elongated tubes, whereas in males, GON-1  
CC is required for migration of a single linker cell to produce a single  
CC elongated tube. The protein is used in the method of the invention. The  
CC specification describes a method for identifying a modulator of a protein  
CC that contains a metalloproteinase domain and a thrombospondin domain. The  
CC method comprises treating a target organism, having a developing gonadal  
CC cell that is responsive to the protein, with a test compound, and  
CC determining any change in migration or shape of the cell attributable to  
CC the test compound. The compound identified are potential therapeutic  
CC modulators of abnormal cell migration and organ shaping, e.g. for  
CC rendering animals (specifically nematodes) sterile and for inhibiting  
CC cancer metastases  
CC  
XX  
SQ Sequence 950 AA;

Query Match 47.8%; Score 2317.5; DB 3; Length 950;  
Best Local Similarity 49.2%; Pred. No. 1,7e-155;  
Matches 450; Conservative 143; Mismatches 242; Indels 79; Gaps 21;

QY 14 LLLLLLLPLIARGAPAPPAAGQASELVPT--RLPG-SAGELALHLSAFKGFVLRLA 70  
DB 21 LLLASITMLLCARGAHGRPTE--EDEELVLPSELRAPGHSTTRRLDAFGQQLHLKIQ 78  
QY 71 PDDSEFLPEFKIERLGGSGRATGER-----GLRGCFSGTVNGEPESLAASLCRGL 123  
DB 79 PDSGFLAPGFTLQTV--GRSPGEAQHLDPTGDLAHCFTSTVNGDPGSAALSLCEGV 135  
QY 124 SGSEFLDGESEFTLPQAGAGSLAQPHRLQMGAPARPLRGPWEVEYEGE---ORQE 179  
DB 136 RGAFYLGGESEFTLPAP-----GVATERLADAVEEESAPRPFHILRRR 181  
QY 180 RGD-----HQDSEESQEEAEGASERPPPLGA-----TSRTKRFVSEA 219  
DB 182 RGSAGACGWDDETLPTSPRPESQNTKQWVPDRPTPODAGKPSGPGSIRKKRFVSSP 241  
QY 220 RFVETLIVADASMAFYGADLQNHILTLMSVAARIYKPSIKNSINLMVYKVLIVDEKW 279  
DB 242 RYVETMLVADQSMADFHSGSLKHLTLFSAARFYKHPISIRNSILVVKILYIEEQK 301  
QY 280 GPEVSDNGGLTLRNFCWQRRFNQPSDRHPHYDTALLTRONFCQEGELCDTLGVADIG 339  
DB 302 GPEVTSNALTLIRNFCWQRRFNQPSDRHPHYDTALLTRQDLCGSH--TCDTLGMADVG 360  
QY 340 TICDPNKSQSVIEDEGLQAHTTLAELGHVLSMHPDSDKPCSTRLFGPMGKHVWAPLFLV 399  
DB 361 TVCDPSRCSVIEDDG-QAFTTAHELGHVNMHPDDAKHCAASLGVSGDSHLMAWSLSS 419  
QY 400 LNOGLPMSPCSAMVITELLDGHDCLIDAFGAALPLPTGLPGMALIYOLDQOCROIFGP 459  
DB 420 LDHSGPMSPGSAVWVTSFLDNGHGECIMDKQNPICKLPSDLPG--TLYDARQCCFTTGE 477  
QY 460 DFRHGNISADQVCAQLWC-HTDGAEPCLCHTKNGSLPWAADGTPCAGPHLCSEGSCLPEE 518  
DB 478 ESKHCPD--AASCTTLMCTGTSGGLVLCCTKH--FPWADTSGEGKWCWCSKCVAKTID 533  
QY 519 VERPKPVVDGGWAPWPGWEGSRTCGGQVQFSHRECKDPEPONGARYCLGRBAKYQSGHT 578  
DB 534 MHHFATPVHSGWPGWPGWGDGSRCTGGGVQYTMRECDNPVKNQKXKCEGRVYRSCNI 593  
QY 579 EECPP-DGKSFREOCCEKKNANYTDMGN--LLQWVPKYAGVSPDRCKLFCARGRSE 635  
DB 594 EDCPPNNGTFRFEEQCEAHNESKASF-GNEPTVEMTPKIAGVSEPKORCKLTCRAKIGY 652  
QY 636 FKVFPAKYIDGTLCPETLALCVRGQCVKAGGCHVVDSPKRLDKGVGCGKANSCKRVSG 695  
DB 653 FVLQPKVVDGTPCGSPDSVSCVQGCYKAGCDRIIDSKKKFKDGVGCGNGSTCKKMSG 712  
QY 696 SLTPNNGYNDIVITPAGATNIDVQKSHPGVQNDGNALALKTADGOYLINGNLASALE 755  
DB 713 IYSTRPGYHDIYITPAGATNIEVHRNQRSGRRNSFLAIRADGTYILNGNFTLSTLE 772  
QY 756 QDILVKGTLIKYSGSIATLERLQSFRLPEPLTVQLTLVPGHVPFPKVKYTFVNDVDF 815

DB 773 QDLTYKGVTVIRYSSSALBERINFSPLKEPITLQVLMV-GHALRPXIKRTYFM----- 825  
QY 816 SMOSSKERATTNIIQPLIHQWYLGWSECSSTCGAGQSRFVECRDPGQASATCKAL 875  
DB 826 ----KKVTESFNALPTF--SEWVIEEWGECSTGSGMQRVVOCRDINGHPASECAKEV 879  
QY 876 KPEDAKPEBSQLCP 889  
DB 880 KPASTRPCADLPCP 893

Search completed: May 7, 2004, 11:49:29  
Job time : 70 secs

Fri May 7 12:18:01 2004

us-09-989-687-4\_1.rat

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:48:10 ; Search time 23 Seconds  
(Without alignments)  
1997.701 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 4853

Sequence: 1 MFPAAPAPRPLFLILLLL.....CNKALKPEDAKPCESQLCP 890

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3919.5	80.8	905	4	US-09-369-364A-9
2	2331.5	48.0	949	4	US-09-568-559-2
3	2328.5	48.0	967	4	US-09-130-491-2
4	2274.5	46.9	481	4	US-09-130-491-8
5	2103.5	43.3	727	4	US-09-445-023A-1
6	2093.5	43.1	727	4	US-09-445-023A-12
7	1842.5	38.0	837	4	US-09-122-126B-2
8	1842.5	38.0	837	4	US-09-634-286A-2
9	1803.5	37.2	608	4	US-09-130-491-13
10	1767.5	36.4	930	4	US-09-369-364A-2
11	1734.5	35.7	930	4	US-09-122-126B-15
12	1734.5	35.7	930	4	US-09-634-286A-15
13	1608	33.1	1882	4	US-09-369-364A-13
14	1562	32.2	874	4	US-09-369-364A-15
15	1476	30.4	2165	4	US-09-800-729-155
16	1460	30.1	551	4	US-09-130-491-16
17	1190	24.5	997	4	US-09-369-364A-7
18	1169	24.1	245	4	US-09-369-364A-11
19	1157	23.8	518	4	US-09-369-364A-22
20	1109	22.9	908	4	US-09-963-791-2
21	1088	22.4	1211	4	US-09-491-522-5
22	1070	22.0	1224	4	US-09-930-872-4
23	1054.5	21.7	757	4	US-09-963-791-24
24	1041	21.5	1205	4	US-09-491-522-11
25	1022.5	21.1	1081	4	US-09-369-364A-17
26	957	19.7	859	4	US-09-369-364A-5
27	680	14.0	589	4	US-09-963-791-12

28	625.5	12.9	438	4	US-09-963-791-22	Sequence 22, Appl
29	549	11.3	566	4	US-09-491-522-7	Sequence 7, Appl
30	495.5	10.2	1745	4	US-09-800-729-89	Sequence 89, Appl
31	492	10.1	507	4	US-09-963-791-10	Sequence 10, Appl
32	488.5	10.1	525	4	US-09-369-364A-21	Sequence 21, Appl
33	437.5	9.0	356	4	US-09-963-791-6	Sequence 6, Appl
34	425.5	8.8	468	4	US-09-963-791-20	Sequence 20, Appl
35	411.5	8.5	491	4	US-09-930-872-2	Sequence 2, Appl
36	406.5	8.4	262	4	US-09-800-729-125	Sequence 125, Appl
37	371	7.6	317	4	US-09-963-791-16	Sequence 16, Appl
38	352	7.3	746	4	US-09-548-797B-4	Sequence 4, Appl
39	336.5	6.9	812	4	US-09-634-098-4	Sequence 4, Appl
40	333	6.9	849	4	US-09-548-797B-6	Sequence 6, Appl
41	329.5	6.8	802	4	US-09-634-098-2	Sequence 2, Appl
42	324	6.7	775	4	US-09-786-256C-15	Sequence 15, Appl
43	324	6.7	775	4	US-09-786-256C-32	Sequence 32, Appl
44	322	6.6	787	4	US-09-548-797B-5	Sequence 5, Appl
45	309.5	6.4	1059	4	US-09-800-729-217	Sequence 217, Appl

# ALIGNMENTS

RESULT 1  
US-09-369-364A-9  
Sequence 9, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hirsakainen, Tina L.  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US/09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 9  
LENGTH: 905  
TYPE: PRT  
ORGANISM: Mus musculus ADAMTS-8  
US-09-369-364A-9

QY	Query Match	80.8%	Score 3919.5;	DB 4;	Length 905;
QY	Best Local Similarity	79.2%	Pred. No. 0;		
QY	Matches 717;	Conservative 74;	Mismatches 99;	Indels 15;	Gaps 4;
QY	1 MFPAAPAPRPLFLILLLL-LPLARGAAPAPAGGASSELVETRLPGSAGELIHLIS 59				
DB	1 MRDPTTGWPLILLLLQLPEPPLVCGAPAGPTGAASELIVETRLPGSASELAFHLS 60				
QY	60 ARKGQFVRLADDSFLAPRFKIERLGSSGATGGERGLRGCFEGTNGGREGSLAAVSL 119				
DB	61 ARGGQFVRLADDSFLAPRFKIERLGSSGATGGERGLRGCFEGTNGGREGSLAAVSL 120				
QY	120 CRLGSGFLDDEEFTIQGAGGSLAOPHRLQWGPAGAR-----PLRPGEW 168				
DB	121 VAGWGSFLIAGEFTIQGAGGSLDQPHRLQWGPAGAR-----PLRPGEW 180				
QY	169 EYETGEGQROERGDHOEDSEESQE--EAGASPPPLGATSTYKRFVSAARVETLL 226				
DB	181 EYEMGQGOERGDHOEDSEESQE--EAGASPPPLGATSTYKRFVSAARVETLL 240				
QY	227 VADASMAAYGADLQNHITLMSVARYKPSIKNSINLWVVKYLIVEDKGEVSDN 286				
DB	241 VADASMAAYGADLQNHITLMSVARYKPSIKNSINLWVVKYLIVEDKGEVSDN 300				
QY	287 GGLTLLNFCNWRFRNOPSDRPHRYDAILLTRONFCQGBELCTLLGYADIGTICDPK 346				
DB	301 GGLTLLNFCNWRFRNOPSDRPHRYDAILLTRONFCQGBELCTLLGYADIGTICDPK 360				
QY	347 SCGVITDESLQAHNLTAHETLGHVLSMPHDDSKPCRLTGRMGKHYMFLFHLNQTLPW 406				

```

Db      361 SCSEVIEDEGLQAAYTLAHELGHYLSMPHDSKPCVRLFGPMGKYHMAFPFIHVKTLFW 420
QY      407 SPSCAMVLTETLDGGHDDCLLDAPGALPLPTGLPGRMALYOLDQOCROIIFGPDPRHCN 466
Db      421 SPSCAYVLTETLDGGHDDCLLDAPGALPLPTGLPGRMALYOLDQOCROIIFGPDPRHCN 480
QY      467 TSAQDVCAQIMC-HTDGAEPILCHTKGSLPMADGTCGCGHLCSEBSCIPSEEVERPKV 525
Db      481 TSVEDICVCLCAHRRDSDEPICHTKGSLMLADGTCGCGHLCIDSCVCLKEDVENPKAV 540
QY      526 VDDGMAFPMGPMGSCSRCTCGGVOFSHRECKDEPFGKGRYCLGRARAKYOSCTHECPD 585
Db      541 VDDGMAFPMGPMGSCSRCTCGGVOFSHRECKDEPFGKGRYCLGRARAKYOSCTHECPD 600
QY      586 KSFRQOCCEKYNAYNTMDGNLLQWVFKYAGVSPDRCKLFCRANGSEFPKFEAKVID 645
Db      601 KSFRQOCCEKYNAYNTMDGNLLQWVFKYAGVSPDRCKLFCRANGSEFPKFEAKVID 660
QY      646 GTLCGPETLACVRGQCVKAGCDHVDPSPKLDKCGVCGKGNCKRKGSLPTVYGN 705
Db      661 GTLCGPETLACVRGQCVKAGCDHVDPSPKLDKCGVCGKGNCKRKGSLPTVYGN 720
QY      706 DIYTPAGATNIDVKQSHPGVQNDGNLYALKTADQYLLNGNLAIISALBODILVKGTL 765
Db      721 DIYTPAGATNIDVKQSHPGVQNDGNLYALKTADQYLLNGNLAIISALBODILVKGTL 780
QY      766 KYSSITLIERLOSFRPLPEPLTVOLLTVPGVEFPKVKYTFEVDVDFSMQSKERAT 825
Db      781 KYSSITLIERLOSFRPLPEPLTVOLLTVPGVEFPKVKYTFEVDVDFSMQSKERAT 840
QY      826 TNIIOPLHAQWVLDGMSGSCSTCGAGMORRYECRDPGQASATCNKALKEDAPCES 885
Db      841 TNIIOPLHAQWVLDGMSGSCSTCGAGMORRYECRDPGQASATCNKALKEDAPCES 900
QY      886 QLCPL 890
Db      901 QPCPL 905

RESULT 2
US-09-568-559-2
; Sequence 2, Application US/09568559
; Patent No. 6649377
; GENERAL INFORMATION:
; APPLICANT: Klonowski, Paul
; APPLICANT: Klonowski, Paul
; APPLICANT: Heller, John
; APPLICANT: Heller, John
; APPLICANT: van Wart, Harold
; TITLE OF INVENTION: Human Aggrecanase and Nucleic Acid
; FILE REFERENCE: ROCH-002
; CURRENT APPLICATION NUMBER: US/09/568, 559
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: 60/133,343
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 949
; TYPE: PRT
; ORGANISM: human
US-09-568-559-2

Query Match      48.0%; Score 2331.5; DB 4; Length 949;
Best Local Similarity 49.8%; Pred. No. 66-189;
Matches 453; Conservative 145; Mismatches 249; Indels 63; Gaps 22;

```

```

QY      11 LPELILLLLLPLARAGAPAPAGQASELVVP--TLPISAGELALHLSAFSGFVLR 68
Db      15 VPTLLALLAALLANS-DLGRPSF--EDELIVLELEAPRH-CTTLRLHAFFQCLDLE 70
QY      69 LAPDSSLAPERKIERLOGSGRATG---ERGIKGFSGTNGEPESLAANSLRCL 123

```

```

Db      71 LRPDSSFLAEGFTILQNV---GRKSGSETPLPENTDLAHCFSGTYNQDPSSAALSLCEGV 127
QY      124 SGSEFLDGEETIOPQAGGSLAOPHRLQRMWGPAGAPLP-----RGEMEV---ET 172
Db      128 RGAFTYLLGEAFVFIQPLPAAS-----ERLATAAEGEKPAELOFLLLRNKGQDVGTGCV 182
QY      173 GEGORQRGHDQDESESESESEBEAEGASBP-----PGLATS--RTKRFVSEARFV 222
Db      183 VDDEPRFTGAETDEDEGEDEGEDEGAQMSBODPALOGVQPTGTSIRKRRFVSHRYV 242
QY      223 ETLIVADASVAAFYADLOHNLITLMSVAARIYKPSIKNSINLMVVKVILVEDEKMGPE 282
Db      243 ETLIVADASVAAFYADLOHNLITLMSVAARIYKPSIKNSINLMVVKVILVEDEKMGPE 302
QY      283 VSDNGGLTIRBFQWQMRFPNPSDRHEHDTALLRKONFCQGEGLCDTLGVADIGTIC 342
Db      303 VTSNPAALTRNFQWQMRFPNPSDRHEHDTALLRKONFCQGEGLCDTLGVADIGTIC 361
QY      343 DPNKSCSVIEDEGLQAAYTLAHELGHYLSMPHDSKPCVRLFGPMGKYHMAFPFIHVLNQ 402
Db      362 DPNKSCSVIEDEGLQAAYTLAHELGHYLSMPHDSKPCVRLFGPMGKYHMAFPFIHVLNQ 421
QY      403 TLPSPGSAVMTETLDGGHDDCLLDAPGALPLPTGLPGRMALYOLDQOCROIIFGPDPR 462
Db      422 TLPSPGSAVMTETLDGGHDDCLLDAPGALPLPTGLPGRMALYOLDQOCROIIFGPDPR 479
QY      463 HCPNTSAQDVCAQIMC-HTDGAEPILCHTKGSLPMADGTCGCGHLCSEBSCIPSEEVER 521
Db      480 HCPNTSAQDVCAQIMC-HTDGAEPILCHTKGSLPMADGTCGCGHLCSEBSCIPSEEVER 535
QY      522 PKPVVDGMAFPMGPMGSCSRCTCGGVOFSHRECKDEPFGKGRYCLGRARAKYOSCTHECPD 581
Db      536 PKPVVDGMAFPMGPMGSCSRCTCGGVOFSHRECKDEPFGKGRYCLGRARAKYOSCTHECPD 595
QY      582 PP-DGSKSFRQOCCEKYNAYNTMD-DGNLLQWVFKYAGVSPDRCKLFCRANGSEFPKFEAKVID 639
Db      596 PP-DGSKSFRQOCCEKYNAYNTMD-DGNLLQWVFKYAGVSPDRCKLFCRANGSEFPKFEAKVID 655
QY      640 EAKVIDGTLGPEETLACVRGQCVKAGCDHVDPSPKLDKCGVCGKGNCKRKGSLPTVYGN 699
Db      656 EAKVIDGTLGPEETLACVRGQCVKAGCDHVDPSPKLDKCGVCGKGNCKRKGSLPTVYGN 715
QY      700 TNYGNDIVTPAGATNIDVKQSHPGVQNDGNLYALKTADQYLLNGNLAIISALBODIL 759
Db      716 AXPGYHDITTPGTATNLEVQKORNGSRNNGSFLAIDAGTYIINDYTLSTLEQDIM 775
QY      760 VKGTLIKTSGSIATLERLOSFRPLPEPLTVOLLTVPGVEFPKVKYTFEVDVDFSMQSK 819
Db      776 YKGVILRYSGSMALEIRISFSPLEKPEPLTVOLLTVPGVEFPKVKYTFEVDVDFSMQSK 824
QY      820 SKERATNIIQPLHAQWVLDGMSGSCSTCGAGMORRYECRDPGQASATCNKALKEDAPCES 879
Db      825 KKKKESFNAI-PTFSA-VWIEWEGSCSKCELGWQRLLVECRDINGQPAASECAKEVRYAS 882
QY      880 AKPCESQOLCP 889
Db      883 TRPCADHPCP 892

RESULT 3
US-09-130-491-2
; Sequence 2, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; PRIOR FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961

```

Fri May 7 12:18:01 2004

us-09-989-687-4\_1.ra1

EARLIER FILING DATE: 1997-08-06  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 967  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-09-130-491-2

Query Match 48.0%; Score 2328.5; DB 4; Length 967;  
Best Local Similarity 49.7%; Pred. No. 1.1e-188;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

11 LFFLLLLLLLLPLRGAPAPRAGGQSELVVP--TTLPSAGEALHLISFGKGYLR 68  
33 VETLLMAAALAAVS-DALGRPS--EDELIVPELEAPGH-GTTRLRIRAFDOQDLE 88  
69 LAPDPSFLAPERKIERLGGSGRATG---ERGLRGCFSGTNGEPELSAALSICGL 123  
89 LRPDSEFLAPGFTLQVW---GRKSGSETPLEPTDLAHCPSGTAVGSDSAALSICGCV 145  
124 SSGFLLDGEFTIQPGAGAGSLAQPRRLQRMGPAPARPP-----RGPEWEV---ET 172  
146 RGAFFYLLGAYFLQPLPAAS-----ERLTAAPGEXPPAPLQFILLRNRQGVGCTGCV 200  
173 GEGQROERGDHODESEESQEEAEAGASEPP-----PPLGATS-RYKRFVSEARFV 222  
201 VDDEPRPTKAEDEDEDETEGDEGPGQWSPODPAIGVGCTGTGSTRKRFVSHRYV 260  
223 ETLIVADASMAAFYADLQNHILTLMSVARIYKPSININMVMKYLIVDEKMP 282  
261 EMTLVADQMAEPHSGKHYLLTLEVSARLYKPSIRKNSLVVVKLVHDKGPE 320  
283 VSDNGELTLRNFQNRFRNOPSDRPHEDTAILLTRQNFQGOBGLCTLGVADIGTIC 342  
321 VTENALTLRNFQNRFRNOPSDRPHEDTAILLTRQNFQGOBGLCTLGVADIGTIC 379  
343 DPNKSCSVIEDGLQAHTLAHELGHVLSMPHDSKPCTRLFGPMKHHWAPLVHLNQ 402  
380 DPNKSCSVIEDGLQAHTLAHELGHVLSMPHDSKPCTRLFGPMKHHWAPLVHLNQ 439  
403 TLPMSCSAVYLTLELDGSHDCLDAPGALPLPTGPGMALYQLDQCRQIGPFR 462  
440 SQMSPCASVMTSFDNGHGBCLMDXQNPLOLEBDDPG--TSVDANRQCFPTGEBCK 497  
498 HCPD--AASITCSITWCTGTSGLVLCQTKH--FPWADGTSGBGKVCINGKCVNKTDRKH 553  
463 HCPNTSAODVCAQLWC-HTDGAEPLOHTKNGSLPMADGTPCGPGLHCSGSLPEEVER 521  
522 PKRVVDGMAFPMGMECESTRTCGGGVQFHSRBECDPEPONGARYCLGRAXYQSGHTEBC 581  
554 FDTPEFGSWGMWMPWDCSRTCGGVQYTRBECNVPKNGGYCGKRVRRSCLEBCC 613  
582 PP-DGKSFEEQCEKYNANYTDM-DGNLLQWPKTAGVSPRDRCLFCRARGRSEFKVF 639  
614 PDNNGKTFREGECEANEFSEKASFGSGPAVEMIPKTAGVSPDRCLICQAGIGFVFL 673  
640 EAKYIDGTLGPEETLAI CVRGQCVKAGCDHVVDSPKLDKGVCGGKNSCRKYSGLTP 699  
674 QPKVYVGTSPSPDSTVCVQGCVKAGCDRIIDSKKFKPCGVCGGNGSGTCKKISGSVTS 733  
700 TNGYNDIVITIPAGATYIDVKSHPGVQNDGNYLAKTADGOYLNGNLAIASIEODIL 759  
734 AKGKHITITIPGATYIEVKQNRGSRNNSFLAIKADGTYILNGYITSTLEQDIM 793  
760 VKGTLTKSGSIATLERLQSFRLPEPLTVQLLTPGEVFPKVKYTFPVNDVDFSMOS 819  
794 YKGVVLYSGSSAALERISFSLKEBLITQVLTV-GNALRPKIKYTFV----- 842  
820 SKERATNIIQPLHQAQWVLGDMSECSGTGAGMORRYVCRPDSGOASATCNKALPED 879  
843 KKKKESFNAL-PTFSA-WVIEBMEGCSKCELMQRLVBCRDINQGPASECKAEKVPAS 900

QY 880 AKPCESOLCP 889  
DB 901 TRPCADHPCP 910

RESULT 4  
US-09-130-491-8  
Sequence 8, Application US/09130491  
Patent No. 6416974  
GENERAL INFORMATION:  
APPLICANT: Holzman, Douglas A.  
APPLICANT: Goodheart, Andrew D.J.  
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
FILE REFERENCE: 09404/041001  
CURRENT APPLICATION NUMBER: US/09/130,491  
CURRENT FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: US 60/054,108  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: US 60/054,961  
EARLIER FILING DATE: 1997-08-06  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 481  
TYPE: PRF  
ORGANISM: Rattus rattus  
US-09-130-491-8

Query Match 46.9%; Score 2274.5; DB 4; Length 481;  
Best Local Similarity 82.1%; Pred. No. 1.5e-184;  
Matches 403; Conservative 35; Mismatches 40; Indels 13; Gaps 2;

401 NOTPMSPGAMTITELLDGSHDCLDAPGALPLPTGPGMALYQLDQCRQIGPFR 460  
3 NKTLPWSPCAVYLTLELDGSHDCLD-----DGHSTYELDQCRQIFGPD 50  
461 FRHCPTSAODVCAQLWC-HTDGAEPLOHTKNGSLPMADGTPCGPGLHCSGSLPEEVER 519  
51 FRHCPTNSVEDICQYLMCRHREDBERICTKNSLMDADGTPCGPGLHCLDGSQVLRREV 110  
520 ERPRVVDGMAFPMGMECESTRTCGGGVQFHSRBECDPEPONGARYCLGRAXYQSGHTE 579  
111 ENPRAVVDGMAFPMGMECESTRTCGGGVQFHSRBECDPEPONGARYCLGRAXYQSGHTE 170  
580 ECPBDGKSFEEQCEKYNANYTDM-DGNLLQWPKTAGVSPRDRCLFCRARGRSEFKVF 639  
171 ECPBDGKSFEEQCEKYNANYTDM-DGNLLQWPKTAGVSPRDRCLFCRARGRSEFKVF 230  
640 EAKYIDGTLGPEETLAI CVRGQCVKAGCDHVVDSPKLDKGVCGGKNSCRKYSGLTP 699  
231 ETKYIDGTLGPEETLAI CVRGQCVKAGCDHVVDSPKLDKGVCGGKNSCRKYSGLTP 290  
700 TNGYNDIVITIPAGATYIDVKSHPGVQNDGNYLAKTADGOYLNGNLAIASIEODIL 759  
291 FSYGNDIVITIPAGATYIDVKSHPGVQNDGNYLAKTADGOYLNGNLAIASIEODIL 350  
760 VKGTLTKSGSIATLERLQSFRLPEPLTVQLLTPGEVFPKVKYTFPVNDVDFSMOS 819  
351 MKGTLTKSGSIATLERLQSFRLPEPLTVQLLTPGEVFPKVKYTFPVNDVDFSMOS 410  
820 SKERATNIIQPLHQAQWVLGDMSECSGTGAGMORRYVCRPDSGOASATCNKALPED 879  
411 SKERATNIIQPLHQAQWVLGDMSECSGTGAGMORRYVCRPDSGOASATCNKALPED 470  
880 AKPCESOLCP 890  
471 AKPCESOLCP 481

RESULT 5  
US-09-445-023A-1  
Sequence 1, Application US/09445023A  
Patent No. 6565988

Page 4

```

DB 649 ASECARKEVKNASIRPCADHFC 6/0

RESULT 6
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6563858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakeozaki, Michinori
; APPLICANT: Ishioke, Keiko
; APPLICANT: Ishida, Yukiko
; APPLICANT: Matsushima, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE OF INVENTION: Composition and method of immunologically analyzing human ADAMTS
; CURRENT FILING DATE: 1999-12-03
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

Query Match 43.1%; Score 2093.5; DB 4; Length 727;
Best Local Similarity 54.0%; Pred. No.6,5e-169; Indels 27; Gaps 12;
Matches 374; Conservative 124; Mismatches 168;

201 EPPPLGATSRTRKRVSEARFVETLLVADASNAAFYGDLONHILTLMSVAARIYKHSI 260
1 EESGP--GSIRKKRRFVSSPRVETMLVADQSVADFHGSLKHAYLLTLFSAARFYGHPSI 58

261 KNSIMLVVTKLIVEDIKMGPEVSDNGILTLNFCWQCRFPQPSDRHREHDTAILLR 320
59 RNSISLIVVVKLIVIEEQGPEVTSNAALLTLNFCWQKHNSPSDRDEHDTAILLFR 118
321 ONFCGEGBLCTLGVADICTICDPNKSQSVIEDEGLQAHTLAHELGHVLSMFDHDSKPC 380
119 QDLCSH--TCDLGLMADVTCVCDPSRSCVIEDDGLQAFTLAHELGHVFNMDHDAKHC 177
381 TRLFGEMKHHVAPLFLVHLNQLTLPWSPGSAAYLTLELDGGRDCLLDARGALLPRTGL 440
178 ASINGVSGDSHMASMLSLDHSQSPSPSAVMVTSFLNGHEBCLMDKFPQNDIKLPSDL 237
441 PERMLYLDOOCROIFGDPFRHCPTNTSADVCAQAWC--HTDGAELPLCHTKNGLPWADG 499
238 PG--TLIDANKQQCFTFGESEKHCSD--AASCTTLMCTGTSGLLVCQTKH--FPWADG 291
500 TPCGPHLCSBESCLPEEVEVERRPVVDGMAWPGMWBCSRTCCGGVQPSHRECKDPER 559
292 TSCGEGKWCYSGKCVKTKDMKFPATPVHSGMWGFWGDCSRTCCGGVQYTMRECDNPVP 351
560 QNGGRVCTLRPAAYOCSHTHEECP--DGKFRBEOGCEKNAVYNTDMDGN--LLQWPKYA 616
352 KNGGKICEGRVARIYRSCNIEDCPDNNKTFREBQCAHNEFSAF--GNBPYEWTPKYA 410
617 GVSFPRDCKLPCFARARSEKVFEEAVYIDGTLGCPETLAIYVGGCYKAGCDHVDSPRK 676
411 GVSFPRDCKLPCFARARSEKVFEEAVYIDGTLGCPETLAIYVGGCYKAGCDHVDSPRK 470
677 LDKCGVGGGKSCRVSGSLPTNGYNDIVTI PAGATVITDVKQESHPEGVNDGYNTIAL 736
471 FDKCGVGGGKSCRVSGSLPTNGYNDIVTI PAGATVITDVKQESHPEGVNDGYNTIAL 530
737 KTDAGVYLLNGNLAIAGIADODILVKGTTILKSGSITLTERLQFRPLPEPLTYQLTLVPG 796
531 RADGVIYINGNTLTLEDDLYTKGTIVARIYSSSSALETIRFSPLEKEPLTYQLVNW--G 589

```

Fri May 7 12:18:01 2004

us-09-989-687-4\_1.raii

Page 5

Cy	79	EVPPEKRYKYFFEPVNDVIDSMOSXKRATNTIIQPLHLAOWVLGPMSSCSSTCGAGQR	856
		:   :   :	
Dd	590	HAIAPRKIKFYFM-----KKKTSEFNAPIPTF--SEWVIEMEGCSTKTCSGWR	637
Cy	857	TVECRDPSQSASATCKALKPEDAKCESQLQP	889
		:   :   :	
Dd	638	VWCQRDLNGHNPASECAKEVFASITRCADLPCEP	670

RESULT 7  
US-09-122-126B-2  
; Sequence 2, Application US/09122126B

```

: APPLICANT: Bristol-Myers Squibb Company
: TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
: FILE REFERENCE: DME909
: CURRENT APPLICATION NUMBER: US/09/122,126B
: CURRENT FILING DATE: 1998-07-24
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2
: LENGTH: 837
: TYPE: prt
: ORGANISM: Homo sapiens
US-09-122-126B-2

```

Query Match	38.0%	Score 1842.5	DB 4	Length 837
Best Local Similarity	45.0%	Pred. No. 1.6e-147		
Matches 373	Conservative 130	Mismatches 244	Indels 85	Gaps 24

```

0Y      5 PAAP-RWIFELILLILL-LPLARKAPAPPAAGGASIVYTRLPGSA-----GELA 55
Db      27 PIVPSWLVWMLILLILLASLLPARLASPPH---REELIVPEPKNGSVLPQSGAPARLL 82
0Y      56 LHLSPGKGFVRLRMPDDSEFLRPEFIEBLSGSGRATGGERGLRGCFPGSYNGEBSIA 115
Db      83 CSLQAFGEFTLLEBEDSGVQVEGLTVQYLQGAPELLGSAE--PGYLTGTINGDEBSYA 140
0Y      116 AVSLCRG-LSGSFLLDGEEFTIQ-----PGAGGSLAOPHRLQWGPAGAPPLPRGPEW 168
Db      141 SLHMOGALLGVLYQNGALHLPQLEGTPNSAGSGGA--HILRRKSPASG---QGPMC 194
0Y      169 EYETGEGQROERGOEDSEBESQEBEACGASERPPLGATS---RTRFVSEARFYVT 224
Db      195 NVXA-----PLGSPSRPRPARFASLSHFVET 222
0Y      225 LVYADASMAFYGADLCNHILTLMSVAARIYKPSIKNSINLMVXVILVEDEKMGPEVS 284
Db      223 LVYADKMAAFHAGAGLKRYYLITMAAAAKFKPSTIRNPVLVTRVLIGSEEBPOVG 282
0Y      285 DNGGLTLRNFQNRPNPDSHPEHYDTAIIILTRONFGGGLDPTLGVAIDIGTICP 344
Db      283 PSAQTLTRSCAWQRGINTPEBDDPHFDYALITRQDLCG-VSTCDLMAADVGVICD 341
0Y      345 NKASVYEDGLOAAHTLAHELGHVLSMPDDSKPCTRELGPWG-KHIVAPLFLVNLQT 403
Db      342 ARSCAIVEDGLOSAFTAAHELGHVFMHLDNKPCISLNGPISTSHVAAFWAHVDP 401
0Y      404 LPWSPGSAMVLTLLDGGHGDCLIDAPGAALPLPTGLPGMALYOLDQCRQILPGDPFRH 463
Db      402 EPWSPSCARFTFLDNGYGHCLIDKKEAPLHPVTPPGKD--YDARQOQLTFGPPSRH 459
0Y      464 CPMTSADQVQAOLWC--HTDGAEPDLCITKXGSLPMAAGTCGPGHLTSESGLPREEVER 521
Db      460 CPQLPPP--CALMGSGHLNG--HAMQTIKS--PMAIGTCGPRQAOMGRGCLHMDLQD 514
0Y      522 PKPVVDGMAFWGFWGECRSRTCGGQVQFSHECKDPEPONGARYCLGRAXYQSCHTEBC 581
Db      515 FNIVQAGMGWGMGWDGCSRTCGGQVQFSSHDCRTPVPRNGKXCEGRKTRFRSCNTEDC 574
0Y      582 PP-DGKSFREBQCKRYAAVNTDMDGNL---LQWVPKXAGVSPDRCLFCRAGRGSEFK 633

```

```

Db      575  PTGSALTPEEBEOCAAYN--HRTDLEKSPFGEMDMVPRKTYGVAHPDQCKLTQCARALGYYY 632
Qy      638  VFEAKYIDSTLCGPETIALCVRGQCVKACDHHVDSPRKLDKCGVCGAKNSCRKATGSL 637
Db      633  VLEPRVYDTPGCSPDSSSCVCGGRCTHAGCRIRIGSKKKPKDCMVCAGDSDGSCSKSGSF 632
Qy      698  TPTNYGNDIVTIPAGATNIDVKQSHSGVONDGNATLAKTADQCYLLNGLNLSAIEQD 757
Db      693  RKPRYGNWVVTIPATHTILVRQGNPG--HRSYIALAKLPDSEVALNGEYTLMPSPTD 750
Qy      758  ILVKGTT-LKYSGLITLERLOSFPPLDEPLTVQLLTPGSEVFPKRYTFEVP 810
Db      751  VILPGAVSLRYSGATASSETLGHQPLAQPLTLQYL-VAGNPQDTRLRKYSFVP 803

```

RESULT 8  
US-09-634-286A-2  
; Sequence 2, Application US/09634286A

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

; SEQ ID NO 2
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-634-286A-2

```

Query Match	38.0%	Score 1842.5	DB 4	Length 837
Best Local Similarity	45.0%	Pred. No. 1.6e-147		
Matches 375	Conservative 130	Mismatches 244	Indels 85	Gaps 24

```

QY 5 PAAB-RMIFLILLILL-LLPLARGAPARRPAAGGASLWVPTRLPGSA-----GELA 55
Db 27 PIVPISWIMWILLILLIASILLPSARLASBP-----REEBIVEPEKINGVLPFGSGAPARLL 82
QY 56 LHLSPRGKPVTKRLAPDDSFLEARPEKIRLIGSGRGATG3BRCLRGCFSGTVNGEPESIA 115
Db 83 CRLQAFGEILLLETFQDSGVQVEGLTVQJLQAPRLDGA--PQTYLTGTINGEPESVA 140
QY 116 AVSLORG-LSGSPILLDGEFFITQ-----POGAGSLAOPHRLQRMWPGAPARPLRGPBW 168
Db 141 SLHMGGALLGLVQYRGAEHLHQPLEGGPNAGGAPGA--HILRRKSSPASG--QGPRC 194
QY 169 EYETGEGROERGHODSEBESQEBEAGABEPPLGAS-----RTRRFVSEARFVYT 224
Db 195 NYTA-----PLGSPSPRRRAKRFASLSRFVET 222
QY 225 LLYADASMAFYGADLONHILLTMSVAARIYKHPISIKSINDMVKVLLVEDEKMWPEVS 264
Db 223 LVAADKMAAFHAGAKRYLLTVMAAAKAFHPSIRNPVSLVTRVLVILGSEEGEPQVG 282
QY 285 DNGGLTLRNPCKMORREFNOPSRHPEHYDTALLTRONFCGEBGLCDTLGVADIGTICP 344
Db 283 PSNAQTLRFSCAWQRGLNTPBEDSDPHFTALLFTRODCG--VSTCDTIGMDVGVICP 341
QY 345 NKSASVITEBEGLOAAHTLAEHGLVLSMHD3KCPTRLFGPMG--KHATMAPLFVHNOT 403
Db 342 ARSCALIVEDGLSATTAHAEHGLVNMHLDNSKPCISLNGPLSTRHYVAPMAHAYDE 401
QY 404 LPMSPSANYLTELDDGHQDCULLDAPGALPLPTGPGRMALYOLDQOCROJFGGDFPH 463
Db 402 EPMSPSARPIITFDLNGYGHCLDREAPFLHPIVTFPKD--YDADRCQQLTFGDSRI 459
QY 464 CPMTSADVNAQIWC--HTDGAEPILCHTNKGLPMDGTPCGGHLCSGSLCPBEVER 521
Db 460 CPQLPBP--CAALWCGSHNG--HAMQTKHS--PMADGTPCGAQAQMGRCILHMQLOD 514

```



Fri May 7 12:18:01 2004

us-09-989-687-4\_1.ra1

Page 6

QY 522 PKRVNDGMAPMGWCEGSRITCGGVQFQSHRECKDEPONGGRYCLGRAXYQSCHTBEC 581  
DB 515 FNIPQAGWGPWGPWGCSTRTCGGVQFQSHRECKDEPONGGRYCLGRAXYQSCHTBEC 574  
QY 582 PP-DGKSFREOQCEKXNANYTMDGNTL--LQWPKYAGVSPDRCKLFCBARGRSEPK 637  
DB 575 PTGSALTFRREOQCAAYN--HRTDLFKSPGPMDMVPRYGVAPQDQCKLTCQARALGYXY 632  
QY 638 VFAKATIDTLCGPBETLALCVRGQCVYAGCDHVDSPRLDKGVCYGGKSGKSRKVSGL 697  
DB 633 VLBPRVVDGTPCSPDSSVCVQGRCHTAGCDRLTIGSKKCFDKCMVCGDGSQSGSKSGSF 692  
QY 698 TPTNYGNDIVTTPAGATNIDVKQSHPGVQNDGNYLAKTADGOYLNGNTLAISAIED 757  
DB 693 RKPRYGNVNTTIPAGATHILVQGNPG--HRSYIALKLPDGSYALNGEYTIMSPFD 750  
QY 758 ILVAGIT-LKXSGIATLERQSRPLPEPLTVQLTVPGEVPPKXTFFP 810  
DB 751 VVLPAGAVSLRYSGATASETLSGHGFLAQLTLQVL-VAGNPQDTRLYSFFVP 803

RESULT 9  
US-09-130-491-13  
Sequence 13, Application US/09130491  
Patent No. 6416974  
GENERAL INFORMATION:  
APPLICANT: Holzman, Douglas A.  
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
FILE REFERENCE: 09404/041001  
CURRENT APPLICATION NUMBER: US/09/130,491  
CURRENT FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: US 60/058,108  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: US 60/054,961  
EARLIER FILING DATE: 1997-08-06  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 13  
LENGTH: 608  
TYPE: PR1  
ORGANISM: Mus musculus  
US-09-130-491-13

Query Match 37.2%; Score 1803.5; DB 4; Length 608;  
Best Local Similarity 52.2%; Pred. No. 2,1e-144; Indels 25; Gaps 12;  
Matches 320; Conservative 113; Mismatches 155;  
QY 271 VLIYDEKMGPEVSDNGGLTLRFNCNMQRRFQSPDRHPEHYDTAILLTREONFCQSGDGLC 330  
DB 4 ILVHDBKXGEVIVSNALLTLRFNCNMQKXNPPSDRAEHYDTAILLTREODLCSQ-TC 62  
QY 331 DTLGVADIGTICDPKSCSVIEDEGLQAHTLAHLGHVLSMHPDSDKPCRTLEFGPMKX 390  
DB 63 DTIGMADVGTCVDPERSCSVIEDGLQAFTTAHLGHVFMPPDDAQCASLGVANQDS 122  
QY 391 HVMALFLVHINOTLPMSPCSAMVLTLLDGGHGDCLDPAALAPLFTGLFGRMALYOLD 450  
DB 123 HVMASMLNSLDSQWSPCSAMVLTSLFDNGHGBELMKXQNPQLPDPDLFG--TSVDAN 180  
QY 451 QOCROIFGDPFRHCNTSAODVCAQWLC-HTDGAPLCHTXGSLPMADGTFCPPGHLS 509  
DB 181 RQCGTTFGBDSXGCPD--AASFCSTLMCTGTSGVLVQYQIKR--FPWADGTSCEBGRKI 236  
QY 510 BEGCLPEEVEVERPKVVDGMAPMGWCEGSRITCGGVQFQSHRECKDEPONGGRYCLGR 569  
DB 237 NKCVCNKIDRKAFDTPFRGSGMGWGPWGCSTRTCGGVQYIMREBDNVPRKXGKXCEGK 296  
QY 570 RAKYQSCHTBECPP-DGKSFREOQCEKXNANYTMD-DGNLLQWVPKYAGVSPDRCKLF 627  
DB 297 RYVYSCNLEDCPDNNGKTFREOQCEAHNEFSKASFGSGPAVEWIPKYAGVSPDRCKLI 356

QY 628 CPARGSEPKYFEAKYVTDGTLCPBETLALCVRGQCVYAGCDHVDSPRLDKXGVCYGGK 687  
DB 357 CQAKGIGYFFFLQPKVVDGTPCSPDSSVCVQGRCHTAGCDRLTIGSKKCFDKCMVCGDGSQSGSKSGSF 416  
QY 688 NSCRKVSGLTPTNYGNDIVTTPAGATNIDVKQSHPGVQNDGNYLAKTADGOYLNG 747  
DB 417 STCKKISGVSATAPGHDIITIPIGAKTNEVQNRGSRNNQSFLATKADQYTLNG 476  
QY 748 NLASAEODILVKTGSLTLEKSGIATLERQSRPLPEPLTVQLTVPGEVPPKXTFFP 807  
DB 477 DYTSLTEODIMYGAVLRYSGSAALEIRISFSPLEPLITVTV-GNALRPKIKTY 535  
QY 808 FVPPNDVPSMGSSEKATNTNIPQLLAQWVLGDMSCSSITCAGMORRYVECDPESG-- 865  
DB 536 FV-----KKKSSFNAL-PTPSA-WVIEBWSGCKTGKGYKKRBLKLSHDG 583  
QY 866 QASATCNKALKE 878  
DB 584 LSHSCDPLKKPK 596

RESULT 10  
US-09-369-364A-2  
Sequence 2, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Appte, Suneel  
APPLICANT: Hirschman, Tina L.  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US/09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 930  
TYPE: PR1  
ORGANISM: mus musculus ADAMTS-5  
US-09-369-364A-2

Query Match 36.4%; Score 1767.5; DB 4; Length 930;  
Best Local Similarity 40.2%; Pred. No. 4.5e-141;  
Matches 380; Conservative 125; Mismatches 355; Indels 85; Gaps 19;  
QY 10 WLPFLILLILL--LLPLARGAFA-----RPAAGGQASLVVP----- 44  
DB 5 WASLILLILLISASCLISLAADSPAAAPADKTRCPQAAAAAEPDQPGEEETREHGLQP 64  
QY 45 -----TRLPQSAEGLALHLTSARGKGFVRLAPDPSFLAPEFKIERLGSGR 90  
DB 65 LAGQRRSGGLVHNDQLSISGSKGVLYVYAGRRFLDLERDVT-VGAAGSLVTAGGGLS 123  
QY 91 ATGGERGLRGCFSGTVNGEPESLAVALCRGLSSFLDGEETTLTQAGAGSLAQPHR 150  
DB 124 ASSGRHG--HCFYRGTVDSFSLAVFDLCGLDGFVAKARVYLIKED--LRGSAEYER 180  
QY 151 LQRMGPAGAR-----PLPRGPEVVEGEQGRORGRHOESEBESQ----- 192  
DB 181 I--YGDGSSRLILHVYNRBGFSEALPPRASCTPASPSPQSPQSPVHRSRRSLAPQL 238  
QY 193 -EEBAEGASEPPPLCATSRTKRFVSEAFVETLLVADASMAFYGADLQNHILTLMSYA 251  
DB 239 LDBSASPSGNGAGPQWMMRRRRRSTSRARQVELLLVADSSMARVYGRGLQHTLLTASIA 298  
QY 252 ARIYKAPSLKNSLNLVAVYVLLVEDEKMPVSDNGULTRVFCNMQRRFQSPDRHPEH 311  
DB 299 NRLYSHASIEHNLRLAVVAVVYVLTDDTSLVSKNAATLKNFCCKHQHQLGDDHEH 358  
QY 312 YDFAILLTRONFCQSGGLCDTLGVADIGTICDPKSCSVIEDEGLQAHTLAHLGHVLS 371  
DB 359 YDAIILFTBEDLGHHS--CDTIGMADVGITCSPERSCAVIEDDGLHAATVVAHEIGHLLG 417

Qy	372	MPHDDSKPCNRLGRLGPMKXKHVMAPLVFNHNGTLPMWSGCSAMLYTELLDGHSGCLLDAPG	431
Db	418	LSHDDSKFCEENFTETEDKRLMSLITSIDASKRWSJCYSATITTEFLDGHGCLLDRLR	477
Qy	432	AALPLPLGKLGQBMALYOLDQOCROIFGPDFRHCNENTSADQVCAQLMCH-TDGAEPILCHTK	490
Db	478	KQIGPEELBQ--TYATQOQCNLTPEEYISVCE---GMDVCARLWCAVVRQGOVCLTK	532
Qy	491	NGSLPMADGTPCGEGLCSBEGSLPEEBEVRPKFVVDGMAFPMGECSTCTGGGYQBS	550
Db	533	--KLPAVEGTPCGKGRVCLQCKVDTKKKYSTS SHGNWGSWGPWGOCSRS CGGGOFA	590
Qy	551	HRECKDEBPONGKRYCLGKRAKYOSCHTEBEPBGKSPRQOCCNKYNAVYTMDSG--NL	608
Db	591	YRHQCNAPRNSGRYCTGKRAIYRSCSVTPCPNGKSPRHHQCCAKNGYO-SDAKGVKTF	649
Qy	609	LOWPKXAGVSPRDRCLFCGARSRSEPKVEAVIIGTLGPEETLAIYRGQCVKACD	668
Db	650	VEMVPEKTAGVLPADVCKLTCAKGTGYVYSPKVTGTETRPSPNSVCTVRGRVYRGCD	709
Qy	669	HVYDSPRKLDKCGVCGGKANSCKRVSGLTPEYNGYNDIYTLPAGATNIDVKQSPHEGQ	728
Db	710	GIISKQYDKCGVCGGNSCTKLIGFNKSKSGYDVRIPGATHIKVRQPKADQT	769
Qy	729	NDGNVLAIKTADGQYLLNGNLAIASHIEDIIVKGTILKYSSSITLRL--QSTRPLPEP	786
Db	770	RFPAYLAKKKTGEYLNGKTYMISTSETIIDINGTVWYSGWSHRDPLHGMGSATKEI	829
Qy	787	LTVQLLVYPGSVFPPK--VKYTFPVPDVEDFSMOSSKERATYNIQPLHAQVYLDWS	843
Db	830	LIVQLATD---LPKALGVRIYSFVPKTKTKQKANSVISHSNKVGPHSTQLQWYTGPM	885
Qy	844	ECSSTCGAGMQRITVECRDPESGOASATCNKALKEBDAKPCESOLC	888
Db	886	ACSRCTDGTGHTRTYQCCDGNRKLAKGCLLSQRPBAFQCLLKKC	930
RESULT 11			
US-09-122-126B-15			
; Sequence 15, Application US/09122126B			
; Patent No. 6451575			
; GENERAL INFORMATION:			
; APPLICANT: Bristol-Myers Squibb Company			
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES			
; FILE REFERENCE: DM6809			
; CURRENT APPLICATION NUMBER: US/09/122,126B			
; CURRENT FILING DATE: 1998-07-24			
; NUMBER OF SEQ ID NOS: 21			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 15			
; LENGTH: 930			
; TYPE: PRF			
; ORGANISM: Homo sapiens			
US-09-122-126B-15			

```

Query Match 35.7%; Score 1734.5; DB 4; Length 930;
Best Local Similarity 39.8%; Pred. No. 2,8e-138;
Matches 378; Conservative 122; Mismatches 348; Indels 101; Gaps 20;

QY 15 LLLLLLLPLNAGAPAPAA---GGG---ASLVLVPTRLPG----- 49
    |||
Db 8 LLLCAFLPLAIVGPATPADKAGCPPTAAAAAPRRQGEVEERAPPEGHPLAOR 67
    |||

QY 50 -----SAGETALHLISAFKGFYLRLPDPSFLAPSEKIERLGGSGRATCG 94
    : : : : :
Db 68 RRSKGLVQNIIDQLYSGGGKGYLVVYAGRRFLDLDEROSTV-----GIMGVYPAAGG 119
    : : : : :

QY 95 E---RGLRGCPFGSGTVNGEPBSLAAVSLCRGLSGSFLIDGEEFTIQOGAGGSLAQPHR 150
    : : : : :
Db 120 TSAPMWRHRSCHFYRGCTVDASPRSLAVFDLCGGIDGFFPAKARVYTLKPL-LRGPAAEEBK 178
    : : : : :

QY 151 LQMGAPAGAPFLIRGPEVEVETGEGROE---RGDHQDESEESQEEZAEAGSEPP--- 203
    : : : : :

```

```

Db 179 GRVYGDSASILL-----HYTREGEFSFEALPPRASCETPASTPEAHEHA.PAHNSPBGRA 232
QY 204 -----PLGATIS-----PRKRVSEARPFETLLVVDASAAAFGAJLONHIL 245
Db 233 ALASQILLDQALSBAGSGSPQJTWRRRRRSTISRARQVELLLVDASAAALYGRBLQHYLL 292
QY 246 TMSVAARIYKHPISIKNSINLMVYKYLIVEDKMKPVSNDGGLTLFNFCNMQRRFNQPS 305
Db 293 TLASIRARLVSHASIENTHIRLAVKVVVLGBDKSELVSXAAATTLXNFCMKQHQHNLQ 352
QY 306 DRPEHYDTAILLTKRONFCQOEGICDPLGVADICGTDPMKSCSVIEDEGLQAAHTLAHE 365
Db 353 DDHEEHYDAALLFRREDLCGHS--CDPLGNADVGTICSPESCAVIEDDGLHAAFTYAE 411
QY 366 LGHTLSPMHDSKECTBLFCPMGKHHTYMAPLFVIANOTLPMWPCSAWYLTBLDGGHDC 425
Db 412 IGHLLGSHSDSKCESTFGSTEDKRLMSSLITSDASKPMKSCTSATITFLLDGHGNC 471
QY 426 LLDAPGAALLPLPTGLPGBMALYOLDQOCROJFGSDPFHCPNTSADQVACIOMCH--TDOAE 484
Db 472 LLDLPKXOILLPEBLPQJ--TYDATOOCNLTFGSEYVCP--GMDVARIKMACAVVRGO 526
QY 485 PLCHTKNGSLPMAOGTPCPGPHLSSEGCLEPEEVEPRKPVAVDGGMAFWGMEGCSRTCG 544
Db 527 MVLCLK--KLVAESTFGCKRIRICLQCKCVDKTKKKYSTSNGMWSGWSGQCSFSCG 584
QY 545 GGVQFSHRECKDPEPONGGRYCLGRAPKYGCSHTBECPPDDKSPREOCCEKNANMYDM 604
Db 585 GGVQFAYHCHNNPARNRNGRYCTGFRALYRCSLMPCPENKSPFHQCCEANQYQ--SDA 643
QY 605 DG--NLLQWVKYAGVSPRDKCKLFCRARGSEEFVFEAKYIDGLCPEPTLAIQVRQC 662
Db 644 KGAVTKFVAMVPRKVASLPAADVCKLTCRAKGGYVYFEPKXTDTEGCRPYNSVCVRGKC 703
QY 663 VKAGCDHVVDSPRKLDKCGVCGGKGNSCRKYSGLTPINYGNDLYVTI.PAGATINIDVOR 722
Db 704 VRTGCDGIGSLQYDKCGVCGGSDNSCTKIVGFENKKSKEGTDVVRLEGGHTLHKVRF 763
QY 723 SHPGVQNDGNAYALKTADGGYTLNMLNLSAIEODILVKGTILKYSJATLERL--OSF 780
Db 764 KAKDQTRTAYALAKKNGEYLINKMWISSETIIDLNGVTMNSGWSHRDPLFHGMGY 823
QY 781 RPLPEPLTVOLI--TVGEVFPPEKXYTFPVENDVDFSMQSSHEKRYTINIIGQLLAAOWYL 839
Db 824 SATKRIILLVQIATDPTK--PLDVAYSFVPEKXSTPKPNASTVSHOSNKGASTTSOPQWY 861
QY 840 GPMSECSSTCGAGWQRATYECDDPSGQASATCNKALKEDEKAPCESOLC 868
Db 862 GPMLASCRITCDGHTTRIVQCQDGNRKLAKGSPQRSAFQCLIKKC 930

```

```

RESULT 12
US-09-634-286A-15
; Sequence 15, Application US/09634286A
; Patent No. 6521436
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLIC PROTEASES
; FILE REFERENCE: DM6909A
; CURRENT APPLICATION NUMBER: US/09/634,286A
; CURRENT FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-634-286A-15

Query Match      35.7%  Score 1734.5;  DB 4;  Length 930;
Best Local Similarity 39.8%  Pred No. 2.8e-138;
Matches 378; Conservative 122; Mismatches 348; Indels 101; Gaps 20;

```

```

QY 15 LLLLLLPLAKGAPAPAA---GGQ---ASELVPTRLPG-----49
DB 8 LLLCAFLPLAAGVAPATPAODKAGOPPTAAAAAQRROGSEEVORAPRPPHPPLAOR 67
QY 50 -----SAGELALHLSAFKGFVLRAPDPSFLAPEKIERLGSGSRATGG 94
DB 68 RREKGLVQNDOLYSGGKGVLYVAGGRFLDLERDSV-----GIAGFVPAGGG 119
QY 95 E---RGLRGCFPSGVNGEPESLAVALSLRGISGFLDGEFTIQPGAGGSLAOPHR 150
DB 120 TSAPWRHSHCFYRGVVDASFRSLAVFDLGGIDGFAYKARVYTLKPL-LGPPAEEBK 178
QY 151 LQRMGAPAGAPLRGPWEVETEGQROE---RGDHOEDSEESQEEAEAGASEPP---203
DB 179 GRVYGGDSARIL-----HYTREGSFEALPRACCEFPAGPREHHAAPHNSNGRA 232
QY 204 -----PGLGATS-----TYKRFVSEARFVETLLVADSMALFYGADLONHL 245
DB 233 ALASQLDQSGALPAGSGSPQTMWRRRRRSISPARQVELLVADSMALFYGGLQHYLL 292
QY 246 TLMSVARIYKHSIKXSINLMVYKYLIVDEKMGPEVSDNGGLTLRFNCNMQARENPS 305
DB 293 TLMSIARLYSHASISNHILAVKVVYLDKDKSLFVSQNAATLTKFCXKHQHNQ 352
QY 306 DRHEHYDTAILLTRQFCGQEGLCDTLGVADITGTCDPNKSCSVTEDEGLAAHTLAHE 365
DB 353 DHEHYDAIILFREDLCGHS--CDTLGMADVGTICSPERSCAVIDEDGLHAFTVAHE 411
QY 366 LGHVLSPHDSKRCCTFLPGMGNHMAFLFVHLNQTLPWSPCSAMVLTLLDGGHDC 425
DB 412 IGHLSHDSKRCCTFLPGMGNHMAFLFVHLNQTLPWSPCSAMVLTLLDGGHDC 471
QY 426 LLDPAALPLPTGLPERMALYQLDQCRQIFGDFHCPNTSAQVCAQLMCH-TDGA 484
DB 472 LLDPRQIILPEBLPGQ--TYDATQCCNLTFGEYVCP---GMVCAALMAVVRQ 526
QY 485 PLCHTKNGLPMADGTCCGPHLCSBESCLPEEVEHPRVVDGMAPMGECERTG 544
DB 527 MYCJTK--KLPAVGTGCGKRIQLQKCVDKTKKTYTSSHONMGSQSGCSBSC 584
QY 545 GGVQFSHRECKDPFONGRKYCLGRRAKYOSCHTECPDPGSEFRECCEKYNAYVTM 604
DB 585 GGVQFAHCKNPRPANNRKYCTGKRAIYSCSLMPCFPNGKSRHQCBAKNGYQ-SDA 643
QY 605 DG--NLQWPKYAVGSPDRCKLFCFARGRSEKYEBAKYIDGTLGPEPTALCYRGCC 662
DB 644 KGVTFEYEWVYKAGVLPADVCKLTCRAKGTGYVVPSPKTYDTECRFPYSNVCVRGKC 703
QY 663 VKAGCDHVDSRKLDKGVCYGGKNSCRKYSGLTFPTYGYNDIVITIPAGATNIDVKOR 722
DB 704 VRTGDDGLTSGKLQIDKCGVCGGNSSTKIVGTFNKSKGYTDVNAIPGATHTIKXRF 763
QY 723 SHPEVQNDGNIALKTADGOYLNGNTAISAIEDDILVKITLKYSGIATLEFL--QSF 780
DB 764 KAKQDRTFAYIALAKKNGEYLINGKWIISTEITIIDINGVWYSGSHRDDFLHGMGY 823
QY 781 RPLRPLLYVOLL--TYRGEVFPKXKTYTFVNDVDFSMQSKERATNTIIOPLHAQVYL 839
DB 824 SATKILLVQIILADPFR--PLDVAYISFFVPEKSTPKYNSVTSHSKNGVSHTSQFQWVT 881
QY 840 GDMSECSSTCGAGMORFTEGCDPSPGASATCNKALKEDEAKPESQIC 888
DB 882 GFWLACSRCTDGTGHTRTVQCDDGRKLAKGCPISQPSAFKQCLLKGC 930

```

```

RESULT 13
US-09-369-364A-13
; Sequence 13, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi

```

```

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ. ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ. ID NO. 13
; LENGTH: 1882
; TYPE: PR
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: MOD. RES
; LOCATION: (468)
; OTHER INFORMATION: Xaa = C
; NAME/KEY: MOD. RES
; LOCATION: (521)
; OTHER INFORMATION: Xaa = Y
US-09-369-364A-13

Query Match 33.1%; Score 1608; DB 4; Length 1882;
Best Local Similarity 37.5%; Pred. No. 4.5e-127; Indels 138; Gaps 24;
Matches 339; Conservative 150; Mismatches 276;

QY 58 LSARFGKGVFLALADDSFLAPEFKIERLGSG-----RATGGERGLRGCFPSGVNGEPE 112
DB 104 LSAFGQGLPFLNLNAGFIAPLFTVTLTGTVNQTKFYSEBAELKHCFYK-----155
QY 113 SLAVALSCRLSGSFLDGEFTIQPGAGGSLAOPH-----RLQRMGP-----156
DB 156 -----RLC-----OYLRAHGRHOPLRNHNKRNHSDKKTKRARKGERINTL 198
QY 157 AG-ARPLRGPEWEVETEGQROERGDHOEDSEESQEEAEAGASEPPPLGATSRTRYF 215
DB 199 AGDVAAALNSGLATEFSAVG-----NKTDMTEKTHR-----RTYRF 236
QY 216 VSEARFVETLLVADSMALFYGADLONHLTLMSVARIYKHSIKXSINLMVYKYLIVE 275
DB 237 LSYRFVFLVADRMVSYHGENLOHYILTLMSIVASIVDPISGLINIVINLVITH 296
QY 276 DEKMEPEVSDNGGLTLRFNCNMQARENPSDRHEHYDTAILLTRQFCGQEGLCDTLGV 335
DB 297 NQDDPSISFNQOTLTKFCQWQHS--NSPG--IHHTAVLLRQDICRAHDKCDTLGL 352
QY 336 ADIGTCDPNKSCSVTEDEGLQAAHTLAHELVHLSNPHDSKRCCTFLPGMGNHMAFL 395
DB 353 AELGTCDPYKSCSISEDSGLSTAFTHAELGHVFNMHDDNNKC--KEEGVKSPOHVNA 411
QY 396 LFVHLNQTLPWSPCSAMVLTLLDGGHDCILLDAP-GAALPLPTGLPERMALYQLDQCR 454
DB 412 TLNFTNPMWMSKSRKTYTEFLDTGCECLINEPSRPYLPVQLPG--ILYNNKQXE 469
QY 455 QIFGDFHCPNTSAQVCAQLMC-HTDGAERLCHTKNGLPMADGTCCGPHLCSBESC 513
DB 470 LIFGSGQYCFPMWQ---CRLMNNVNGVNHKGRCTQH--PFWAGTECEPEKHKXGFC 524
QY 514 LBEVEYERKPYVDGMAPMGPMGECSTCCGGYQFHSRECKDPFONGRKYCLGRRAKY 573
DB 525 VPR--EMDVPTVDSGWSMSPFGTCSRCCGGIGITFALRECNRPKPKAGKCYVRBMKF 581
QY 574 QSCHTECPDPGSEFRECCEKYNAYVTMDGNT--LQWPKYAVGSPDRCKLFCFAR 631
DB 582 KSNTPRCLKQKRDPRDECAHFDGHP-NINGLLPNYRWVVKYSGILMKDKRCKLFCFAR 640
QY 632 GNSFVFEFAKYIDGTLGPEPTALCYRGCCYKAKGCDHVDSRKLDKGVCYGGNSRCK 691
DB 641 GNTAVYQLDRVYIDGTPCGQDNDICVQGLCSQACGDHVLNSKARDKCGVCGDNSCK 700
QY 692 KYSGSLPTNYGYNDIVITIPAGATNIDVKORSHPEVQNDGNIALKTADGOYLNGNTAI 751
DB 701 TYAGTNTYHIGNVYLRIPAGATNIDVQHSFSGETDDNYALSSKGEFLNGNFV 760
QY 752 SAIEDDILVKITLKYSGIATLERLOSFRPLRPLVOLLVVPGEVFPKXKTYTFVFN 811

```

Db 761 TNAKEIRIGNAVYSGSETAVERINSTDRIEGELLLOVLSV-GKLYNPVYRSGFNPI 819  
Qy 812 D-----VDFSMQSKERAT-----NTIQPL- 832  
Db 820 EDKPOQFYNNSHGPMWACSKPCQGERKRLVCTRESQDQTVSDQRCDLRPFQGHTEPCG 879  
Qy 833 --LHAQWVLGDMSECGSTGAGWQRTVRCRDPSCGASAT-----CNKALKEPADKPC 883  
Db 880 TGGDLHMVAVSRSECSAQCGLGRTLDYCAKYSRLDGTEKVDGFCSSHPKPSNRXC 939  
Qy 884 ESQ 886  
Db 940 SGE 942  
  
RESULT 14  
US-09-369-364A-15  
; Sequence 15; Application US/09369364A  
; Patent No. 6391610  
; GENERAL INFORMATION:  
; APPLICANT: Apte, Suneel  
; APPLICANT: Hurskainen, Tiina L.  
; APPLICANT: Hirohata, Satoshi  
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
; FILE REFERENCE: 26473/4007/10-30-00  
; CURRENT APPLICATION NUMBER: US/09/369,364A  
; CURRENT FILING DATE: 1999-08-06  
; NUMBER OF SEQ. ID NOS: 31  
; SOFTWARE: Patentn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 874  
; TYPE: PRT  
; ORGANISM: Mus musculus ADAMTS-9  
US-09-369-364A-15  
  
Query Match 32.2%; Score 1562; DB 4; Length 874;  
Best Local Similarity 37.3%; Pred. No. 1,1e-123;  
Matches 322; Conservative 152; Mismatches 266; Indels 124; Gaps 22;  
  
Qy 115 AAVSICGLSGSFLDGEFTIQP-----QAGGSLAOPHRLQWGPAGAPLPFGPPM 168  
Db 3 AVISLCGGMGTFRSHDGYFIEPLQSVDBQDEBEQNKPHIYRST-----PGR 53  
Qy 169 EVETGE--GQROERGHQSDS-----EEESQEEKREGASEPPPLGAT----- 209  
Db 54 EPSTGKACATSELNKSHSKDKIRMRKRRKNSIADVDALKSGIATKVLGSGYNQTN 113  
Qy 210 -----SPTKRVSAAREVELTLVADASMAAFYGADLQNHITLMSVAA-YYKHPSTX 261  
Db 114 NTRDRMNRKTKRLSLPRFVEVNVADHRMVLHGKLNLOHLYITLMSIYKSSIG 173  
Qy 262 NSINIMVVKVLIIVDEKMGPEVSDNGGLTLRNFCMWRPNRPNQSDRPHY-----DTA 315  
Db 174 NLIVIVLVNIVNEGEPYINFMATLKNFCQW-----HSKNYLGIOHDTA 224  
Qy 316 ILTLRQNFQCGEGICDTLGVADITICDPNKSQSVIIDEQLQAHLTAHLGLVLMSPH 375  
Db 225 VLVTRREDICRAQDCKDTLGLAELGTICDPYRSCSISDSGLSTAFTHAHLGVFNPHD 284  
Qy 376 DSKCCTRLFGPMGKHVWAPLFLVHINQTLPMSPCSAYITLTDLGGHDCILDPAG-AL 434  
Db 285 DSNKC-KEGVKSPQHVAAPLNTFYTPMWSKSRKXITEPLDTGYGECILNEPASRTY 343  
Qy 435 PLPTGLPGRVALYLDQCCROIIFPPDRHCPNTSAQVCAQWNC-HTDGAEPLCHTKNGS 493  
Db 344 PLPQOLPQ--LLYVNNKQCELIIFGSGQVCEYMMQ-----CRLMNNVDAHKGCKQH-- 396  
Qy 494 LPMADGTCGPGHLCSEGSCLPREEVERPKPVVDGMAFPMGPWECRSTGGGVQFSHRE 553  
Db 397 TPMADGTCGPGHLCSEGSCLPREEVERPKPVVDGMAFPMGPWECRSTGGGVQFSHRE 453  
Qy 554 CKDEPONGARYCLGRRAKYOSCHTEBCPPDGKSFREOCCCKTAAVYTMDEL--LQW 611

Db 454 CNRPEPNNQGVYGVRRMKFSNTEPCKMKQKDFREECQAFHDPKHF-NINGLLPSVRW 512  
Qy 612 VPKYAGVSPDRCLTPRARGSEFFKVEAKVLDGTLGSEETLACVGGCVVAGCDHYV 671  
Db 513 FPKYSGILMDRCLTPRVAAGNTAYVQLRDRVIDGTPCGQDTNDICVQGLCRQGDHIL 572  
Qy 672 DSPKRLDKCGVCGGKSGKRSKVSGLTPYNYNDIYTPAGATNIDVKORSHPGVQNDG 731  
Db 573 NSKTRKCKGICGGSDNSCKTVAQTFTVYGVNTVIRIAGATSIDVRQHSFGKSED 632  
Qy 732 NYLAKTADQVYLINGNLALSAIEQDILVGTIKYSGSLATLERLQSPPLPEPLTVOL 791  
Db 633 NYLALNSKGEFFLNGPFVYMSKEEVRVSAVLEYSQDNVCERLNCTDRIBELLQV 692  
Qy 792 LTVGEVPPKRYTTFVNDVDPSMQ----- 818  
Db 693 LSV-GKLYNPVYRSGFNPIE-DKPOQFYNNSHGPMWACSKPCQGERPKLVCTRESQD 750  
Qy 819 SKKERATNTIQ--PLIHA-----QWLDGMSGSGTGAGWQRTVRCRDPSCGASA 869  
Db 751 TVSDQRCDLRFPQBPVTEACGTDCDLRMVAVSRSECSAQCGLGRTLDYCAKYSRMDGK 810  
Qy 870 T-----CNKALKEPADKPCESQ 886  
Db 811 TEKYDSEFSCSPRPSNOEKCSGE 834  
  
RESULT 15  
US-09-800-729-155  
; Sequence 155; Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ. ID NOS: 217  
; SOFTWARE: Patentn Ver. 2.0  
; SEQ ID NO 155  
; LENGTH: 2165  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-155  
  
Query Match 30.4%; Score 1476; DB 4; Length 2165;  
Best Local Similarity 34.1%; Pred. No. 8,9e-116;  
Matches 327; Conservative 150; Mismatches 319; Indels 164; Gaps 29;  
  
Qy 46 RLPGSAGEIALHLSAFGKGFVRLAPDPSFL-----APEFKIRLQGS 88  
Db 96 RLQVARDQ-----GHACHRLRSDDAVYIVLHRMNOIPDSHNSKVPFHSNSNFAFM 148  
Qy 89 GKAATGEGKLG-----GFSGTVNGEPELSAAVSLC---RLSGSFLDGEFTIQ 137  
Db 149 VLYLDESEBVRGMSKRTPDCLYRAHVGKHQ-SIVLDCSEBDLYKMLAPSGIHIVE 207  
Qy 138 P-----QAGGSLAOPHRLQWGPAGAPLPFGPEWETEGSGQROERGDH-OEDSEE 185  
Db 208 PIISGNGTEHDASHRHQHLVAKFDPMHKSF-----DHINSTSVN 248  
Qy 190 ESQEEBESASRPPPLGATSRKTFV-SBARFVETLIVADASMAAFYGADLQNHITLTM 248  
Db 249 ETEITVAIQDQMEVIERKARSRRAANSMDHYEVVLYADIKYEHGRSLIEDVLTIF 308  
Qy 249 SVAAIRYKPSIKNSINIMVVKVLIIVDEKMGPEVSDNGGLTLRNFCMWRPNRPNQSDR 308  
Db 309 STVASIYRQSIKASINVVVVLIVLKTENAGPRITONAAOQTLQDFCMQOQYNDPDDSS 368

Fri May 7 12:18:01 2004

us-09-989-687-4\_1.ra1

Page 10

QY 309 PEHYDPAILLTRONFCQOEGLCDTLGVADIGTICDPNKSCSVTEDEGLQAAHTLAHELGH 368  
DB 369 VOHHDDVALILTRKIDICRSQKCDTLGLAELGTMCQKSCALLIEDNGLSAFTIAHELGH 428  
QY 369 VLSMPHDDSKPCSTRLPQPMGK-----HHVAPLFWHLNQTLPMSPCSA 411  
DB 429 VFSIPHDDEKCS-TYMPVNVKVCFKQSTKFDKTOFQNNFHIMAPLTLEYNTHPMSWSPCSA 487  
QY 412 MYTELLDGGHG--DCLLDAPGAAL----PLFTGLPGRMALYQLDQOCROIFEGPDRHCP 465  
DB 488 GMLERFLENNRGOTQCLFDQVERRYEDVFVRDEPRK--YDAHQCKVFBGPAELCP 545  
QY 466 NTSADVCACQIMCHT-DGAPLCHTKNGSLPMADGTPCGPGH--LCSGSL--PEEV 519  
DB 546 ---YMPGCRRLMCAITFGSQMGCTGH--MPNADGTPCDESRSMFCHHGACVRLAPESLT 600  
QY 520 ERFPVVDGMAFPMGEGSCRTGGVQSHRECKDPEPONGRYCLGRAYQSCHE 579  
DB 601 K-----IDGQWDRSWGECSTGCGVQKGLRDCSPKPRNGKTCVGRERYSCNTQ 655  
QY 580 ECPDGKSFRCQCEKTN--AYNYTDMGNLQWPKYAGVSPDRCKLFCRARGSEFK 637  
DB 656 ECPMDTQPYREVQSEFNKDIGQVASTNTHTWPKYANVAPNERCKLYCRLSGAIFY 715  
QY 638 VFEAKVIDGTLCGETLAICVRGQCVAGCDHVVDSPKLDKCGVCGKGNCKRYSGSL 697  
DB 716 LTRDKVVDGTPCDRNGDDICVAGACMPAGCDHOLHSTRDKCGVCGGDDSCKVKGT 775  
QY 698 TPT-NGVNDIVTIPAGATNDVKQBSHPGVONDGNVLAUKTADGOYLNGNLATSAIBQ 756  
DB 776 NEQGTFGYNEWKIPASANTIDIKQKYNMKEDDNYLSRAANGELNGHFOVSLARQ 835  
QY 757 DILVGTILKXSGSIATLERLOSFRPLPEPLTVQLLTVGSEVPPKVKYTF--VPNDY 813  
DB 836 QIAFQDVTLEYSGSDAIIERINGTPIRSDIYVHLSVSGH--PPDISYEMTAAPNAV 893  
QY 814 -----DFSM-QSKERATNIIOL----- 832  
DB 894 IRPISALYIMRVTDITTECDRACRGOQSOXLMCLMSTHQSHDRVCQNVLPKQATRM 953  
QY 833 ---LHAQWVLGDNSECSSTGAGMORRTVECRDPSC--QASAT--CNKALKEBDAKPC 883  
DB 954 CNIDOSTMITEDEVSSGAKCGSGQKQRVSCVMEGBRQTPASEHLCDRNSKPSDIASC 1013

Search completed: May 7, 2004, 11:52:17  
Job time : 26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:57:22; Search time 23 Seconds  
(without alignments)  
1997.701 Million cell updates/sec

Title: US-09-989-687-4  
Perfect score: 890  
Sequence: 1 MFPAAPAPWLPFLILL.....CNKALKEPDAPKESQGLPL 890

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 389414 segs, 51625971 residues

Word size: 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued Parents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A-COMB.pep.\*  
2: /cgn2\_6/prodata/2/1aa/5E-COMB.pep.\*  
3: /cgn2\_6/prodata/2/1aa/6A-COMB.pep.\*  
4: /cgn2\_6/prodata/2/1aa/6B-COMB.pep.\*  
5: /cgn2\_6/prodata/2/1aa/PCITUS-COMB.pep.\*  
6: /cgn2\_6/prodata/2/1aa/backfillse1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	24.4	245	US-09-369-364A-11	Sequence 11, Appl
2	35	3.9	905	US-09-369-364A-9	Sequence 9, Appl
3	30	3.4	481	US-09-130-491-8	Sequence 8, Appl
4	13	1.5	2165	US-09-800-729-155	Sequence 155, App
5	12	1.3	438	US-09-963-791-22	Sequence 22, Appl
6	12	1.3	551	US-09-130-491-16	Sequence 16, Appl
7	12	1.3	589	US-09-963-791-12	Sequence 12, Appl
8	12	1.3	727	US-09-445-023A-12	Sequence 12, Appl
9	12	1.3	757	US-09-963-791-24	Sequence 24, Appl
10	12	1.3	837	US-09-122-126B-2	Sequence 2, Appl
11	12	1.3	837	US-09-634-286A-2	Sequence 2, Appl
12	12	1.3	859	US-09-369-364A-5	Sequence 5, Appl
13	12	1.3	908	US-09-963-791-2	Sequence 2, Appl
14	11	1.2	263	US-08-300-903A-2	Sequence 2, Appl
15	11	1.2	263	US-08-988-197-2	Sequence 2, Appl
16	11	1.2	655	US-08-148-910-12	Sequence 12, Appl
17	11	1.2	655	US-08-448-937A-12	Sequence 12, Appl
18	10	1.1	58	US-09-800-729-168	Sequence 168, App
19	10	1.1	205	US-09-800-729-149	Sequence 149, App
20	10	1.1	207	US-09-800-729-113	Sequence 113, App
21	10	1.1	208	US-09-800-729-151	Sequence 151, App
22	10	1.1	231	US-08-220-379B-7	Sequence 7, Appl
23	10	1.1	231	US-08-243-545-2	Sequence 2, Appl
24	10	1.1	231	US-08-993-962-2	Sequence 2, Appl
25	10	1.1	231	US-09-160-841-2	Sequence 2, Appl
26	10	1.1	231	US-08-669-692-2	Sequence 2, Appl
27	10	1.1	231	US-08-444-626-2	Sequence 2, Appl

28	10	1.1	231	5	PCT-US94-05365-2	Sequence 2, Appl
29	10	1.1	231	5	PCT-US95-03866-6	Sequence 6, Appl
30	10	1.1	467	4	US-09-148-545-134	Sequence 134, App
31	10	1.1	467	4	US-09-907-794A-195	Sequence 195, App
32	10	1.1	467	4	US-09-905-125A-195	Sequence 195, App
33	10	1.1	467	4	US-09-902-775A-195	Sequence 195, App
34	10	1.1	480	2	US-08-828-488-8	Sequence 8, Appl
35	10	1.1	480	4	US-09-239-689A-8	Sequence 8, Appl
36	10	1.1	480	4	US-09-702-705-336	Sequence 336, App
37	10	1.1	480	4	US-09-736-457-336	Sequence 336, App
38	10	1.1	480	4	US-09-614-124B-336	Sequence 336, App
39	10	1.1	480	4	US-09-671-325-336	Sequence 336, App
40	10	1.1	480	4	US-09-589-184-336	Sequence 336, App
41	10	1.1	492	4	US-07-794-393-4	Sequence 4, Appl
42	10	1.1	492	1	US-08-001-711-4	Sequence 1, Appl
43	10	1.1	514	4	US-09-800-729-124	Sequence 124, App
44	10	1.1	608	4	US-09-130-491-13	Sequence 13, Appl
45	10	1.1	633	4	US-09-919-060-13	Sequence 13, Appl

#### ALIGNMENTS

RESULT 1  
US-09-369-364A-11  
Sequence 11, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Aptec, Suneel  
APPLICANT: Hurskainen, Tiina L.  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US/09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 245  
TYPE: PRT  
ORGANISM: Homo sapiens ADAMTS-8  
US-09-369-364A-11

Query Match 24.4%; Score 217; DB 4; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.8e+203;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 AEGASEPPPIGATSRTRGVSEARFVETLLVADASNAAFGADLQNHITLMSVAARIY 255  
DB 2 AEGASEPPPIGATSRTRGVSEARFVETLLVADASNAAFGADLQNHITLMSVAARIY 61  
QY 256 KHPSTKNSINIMVYVLLVEDEKMGPEVSDNGGTLIRFCNMORFNPQSRHPHYDTA 315  
DB 62 KHPSTKNSINIMVYVLLVEDEKMGPEVSDNGGTLIRFCNMORFNPQSRHPHYDTA 121  
QY 316 ILLTRNQFCGEGICDITLVADIGTICDPNKSQSVIEDEGLQAHTLAHEIGTVLSMHPD 375  
DB 122 ILLTRNQFCGEGICDITLVADIGTICDPNKSQSVIEDEGLQAHTLAHEIGTVLSMHPD 181  
QY 376 DSKPCTRLFGPMGKHVNAPIFVHINQTLPMSPCSAM 412  
DB 182 DSKPCTRLFGPMGKHVNAPIFVHINQTLPMSPCSAM 218  
RESULT 2  
US-09-369-364A-9  
Sequence 9, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Aptec, Suneel  
APPLICANT: Hurskainen, Tiina L.  
APPLICANT: Hurskainen, Satsoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US-09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 905  
TYPE: PRT  
ORGANISM: Mus musculus ADAMTS-8  
US-09-369-364A-9

Query Match	3.9%	Score 35	DB 4	Length 905
Best Local Similarity	100.0%	Prod. No.	2.9e-25	
Matches	35	Conservative	0	Indels 0
		Mismatches	0	Gaps 0

Qy 617 GVSFPRCKLFRARGSEFKYFEAKVIDGTLGP 655  
Db 632 GVSFPRCKLFRARGSEFKYFEAKVIDGTLGP 666

```

RESULT 3
US-09-130-491-8
; Sequence 8, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: HOLTMEN, Douglas A.
; APPLICANT: Goodheart, Andrew D.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-09-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-8

```

Query Match	3.4%	Score 30	DB 4	Length 481
Best Local Similarity	100.0%	Pred. NO.	1.2e-20	
Matches 30	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```
QY      702 YGNDIVTIPAGATNIDVKQRSHPGVQNDG 731
          |||||
DB      293 YGNDIVTIPAGATNIDVKQRSHPGVQNDG 322
```

```

RESULT 4
US-09-800-729-155
; Sequence 155, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.,
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044E1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 155
; LENGTH: 2165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-155

```

```
Query Match      1.5%; Score 13; DB 4; Length 2165;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 536 WGECSRTC GGGVQ 548  
Db 612 WGECSRTC GGGVQ 624

```

RESULT 5
US-09-963-791-22
: Sequence 22: Application US/09963791
: Patent No. 6649399
: GENERAL INFORMATION:
: APPLICANT: Donoho, Gregory
: APPLICANT: Turner, C. Alexander Jr.
: APPLICANT: Friedrich, Glenn
: APPLICANT: Scoville, John
: APPLICANT: Zambrowicz, Brian
: APPLICANT: Sands, Arthur T
: TITLE OF INVENTION: No. 6649399e1 Human Proteases and Polynucleotides Encoding the S
: FILE REFERENCE: LEX-0105-USA
: CURRENT APPLICATION NUMBER: US/09/963,791
: CURRENT FILING DATE: 2000-12-08
: PRIOR APPLICATION NUMBER: US 60/169,769
: PRIOR FILING DATE: 1999-12-09
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 22
: LENGTH: 438
: TYPE: prt
: ORGANISM: Homo sapiens
US-09-963-791-22

```

```

Query Match      1.3%;   Score 12;   DB 4;   Length 438;
Best Local Similarity 100.0%;   Pred. No. 0.0043;
Matches 12;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0

```

Qy	536	WGEC	SRT	CGG	GV	547
Db	416	WGEC	SRT	CGG	GV	427

```

RESULT 6
US-09-130-491-16
; Sequence 16, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodrich, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 551
; TYPE: FRT
; ORGANISM: Rattus rattus
US-09-130-491-16

```

Query Match	1.3%	Score 12;	DB 4;	Length 551;
Best Local Similarity	100.0%;	Pred.No. 0.0053;		
Matches 12; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY 706 DIVIPAGATNI 717



DB 324 DIVTIPAGANI 335

## RESULT 7

```
US-09-963-791-12
; Sequence 12, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedlich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399e1 Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 589
; SEQ ID NO 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-12
```

## Query Match

Best Local Similarity 1.3%; Score 12; DB 4; Length 589;  
Best Local Similarity 100.0%; Pred. No. 0.0056;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTGGGV 547

DB 567 WGECSRTGGGV 578

## RESULT 8

```
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Ei-ji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishida, Keiko
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: 057092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12
```

## Query Match

Best Local Similarity 1.3%; Score 12; DB 4; Length 727;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 DIVTIPAGANI 717

DB 500 DIVTIPAGANI 511

## RESULT 9

```
US-09-963-791-24
; Sequence 24, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedlich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399e1 Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 757
; SEQ ID NO 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-24
```

## Query Match

Best Local Similarity 1.3%; Score 12; DB 4; Length 757;  
Best Local Similarity 100.0%; Pred. No. 0.007;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTGGGV 547

DB 416 WGECSRTGGGV 427

## RESULT 10

```
US-09-122-126B-2
; Sequence 2, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-122-126B-2
```

## Query Match

Best Local Similarity 1.3%; Score 12; DB 4; Length 837;  
Best Local Similarity 100.0%; Pred. No. 0.0077;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTGGGVQPS 550

DB 532 CSRTGGGVQPS 543

## RESULT 11

```
US-09-634-286A-2
; Sequence 2, Application US/09634286A
; Patent No. 6521436
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909A
; CURRENT APPLICATION NUMBER: US/09/634,286A
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
US-09-634-286A-2
```

```

; SEQ ID NO 2
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-634-286A-2

Query Match          1.3%; Score 12; DB 4; Length 837;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGCGV 550
DB 532 CSRTCGGCGV 543

RESULT 12
US-09-369-364A-5
; Sequence 5, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apple, Suneeel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-6
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (450)
; OTHER INFORMATION: Xaa = L
US-09-369-364A-5

Query Match          1.3%; Score 12; DB 4; Length 859;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGCV 547
DB 519 WGECSRTCGGCV 530

RESULT 13
US-09-963-791-2
; Sequence 2, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 664939991 Human Proteases and Polynucleotides Encoding the SA
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 908
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-2
```

```

Query Match          1.3%; Score 12; DB 4; Length 908;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGCV 547
DB 567 WGECSRTCGGCV 578

RESULT 14
US-08-300-903A-2
; Sequence 2, Application US/08300903A
; Patent No. 5581630
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Gail, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,903A
; FILING DATE: 06-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/236,919
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-300-903A-2

Query Match          1.2%; Score 11; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLLL 24
DB 18 LLLLLLLLLL 28

RESULT 15
US-08-988-197-2
; Sequence 2, Application US/08988197
; Patent No. 6548065
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Gail, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
```

Fri May 7 12:17:59 2004

us-09-989-687-4.ra1

Page 5

STREET: 51 University Street  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple Operating System 7.1  
SOFTWARE: Microsoft Word for Apple, Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,197  
FILING DATE:  
CLASSIFICATION: 121097  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/300,903  
FILING DATE: 06-SEPTEMBER-1984  
APPLICATION NUMBER: USN 08/236,919  
FILING DATE: 06-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2822-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-587-0430  
TELEFAX: 206-233-0644  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-988-197-2

Query Match 1.2%; Score 11; DB 4; Length 263;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 14 LLLLLLLLLL 24  
Db 18 LLLLLLLLLL 28

Search completed: May 7, 2004, 12:01:23  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:51:46 ; Search time 65 Seconds  
(without alignments)  
3868.727 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890  
Sequence: 1 MFPAAPAPRMWLPFLILLT.....CNKALKPEDAKPCESQLCPL 890

Scoring table: OLIGO  
Gapop 60.0 , Gapept 60.0

Searched: 1586107 seqs, 282547505 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: A\_Geneseq29Jan04:\*

1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Prod. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	890	100.0	890	2	AAV49502 Human MET
2	890	100.0	890	4	AAV49502 Human MET
3	675	75.8	890	6	ABP96306 Human ADA
4	649	72.9	889	4	AAV4946 Human ADA
5	362	40.7	680	3	AAV21251 Human met
6	263	29.6	364	4	AAV09710 Human gen
7	263	29.6	364	7	ADCC2088 Human pro
8	217	24.4	245	4	AAV72285 Human ADA
9	35	3.9	905	4	AAV72284 Human ADA
10	30	3.4	481	2	AAV04145 Human ADA
11	24	2.7	98	3	AAV21262 Rat meta
12	13	1.5	2150	3	AAV53898 Human sec
13	13	1.5	2150	3	AAV53898 Human sec
14	12	1.3	18	1	AAV81021 Sequence
15	12	1.3	19	4	AAV81021 Sequence
16	12	1.3	20	4	AAV81021 Sequence
17	12	1.3	54	3	AAV08135 Human aci
18	12	1.3	438	4	AAV03582 Human pro
19	12	1.3	575	7	ADV85489 Human ags
20	12	1.3	589	4	AAV03577 Human ags
21	12	1.3	625	5	AAV48394 Rat aggre
22	12	1.3	757	4	AAV03583 Human pro
23	12	1.3	837	2	AAV75425 Human ags
24	12	1.3	837	3	AAV9429 Human ags
25	12	1.3	837	4	AAV78228 Human ags

## ALIGNMENTS

RESULT 1	AAV49502	standard; protein: 890 AA.
ID	AAV49502	
XX	AAV49502;	
XX	10-JAN-2000	(first entry)
DT	10-JAN-2000	(first entry)
DE	Human METH2 protein.	
XX	Human METH2 protein.	
XX	Human METH2 protein.	
KW	Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;	
KW	cancer; diagnosis; hyperproliferative disorder; autoimmune disease;	
KW	angiogenesis inhibitor; abnormal wound healing; inflammation;	
KW	rheumatoid arthritis; psoriasis; endometrial bleeding disorder;	
KW	diabetic retinopathy; macula degeneration; haemangioma; detection;	
KW	arterial-venous malformation; immune deficiency.	
OS	Homo sapiens.	
XX	Homo sapiens.	
XX	MO9937660-A1.	
XX	29-JUL-1999.	
XX	29-JUL-1999.	
XX	22-JAN-1999;	99WC-US001313.
XX	23-JAN-1998;	98US-0072298P.
XX	28-AUG-1998;	98US-0098539P.
XX	(IRUE/)	IRUELA-ARISPE L.
XX	(HAST/)	HASTINGS G A.
XX	(RUBS/)	RUBEN S M.
XX	IrueLA-Arispe L,	Hastings GA, Ruben SM,
XX	WPI; 1999-590684/50.	
XX	DR N-PSDB; AA232001.	
XX	New isolated metalloprotease thrombospondin polypeptides, useful for	
XX	treating hyperproliferative disorders, cancers or autoimmune disorders.	
XX	Claim 10; Fig 2; 457p; English.	
XX	AA232000 and AA232001 encode, and AAV49501 and AAV49502 represent, human	
XX	metalloprotease thrombospondin (METH) proteins METH1 and METH2	
XX	respectively. METH1 and METH2 have been found to be potent inhibitors of	
XX	angiogenesis both in vitro and in vivo. They can be used for treating	
XX	cancer and other disorders related to angiogenesis including abnormal	
XX	wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial	

bleeding disorders, diabetic retinopathy, some forms of macula degeneration, haemangiomas, and arterial-venous malformations. They may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilisation (chemotaxis) of immune cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or infectious. They can also be used to treat inflammatory conditions, both chronic and acute conditions. The products can also be used for detection and diagnosis. AA322002 to AA322080, and AA149503 to AA149511 represent sequences given in the exemplification of the present invention

SQ Sequence 890 AA;

Query Match	Score	DB 2	Length
100.0%	890	890	890

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 890; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0.

[illegible]

Dy 841 DMSGSSSTCGAGMGRRTVECRDEPSGQASATCNALKEPAKPCESQLCPL 890

Dd 841 DMSGSSSTCGAGMGRRTVECRDEPSGQASATCNALKEPAKPCESQLCPL 890

## RESULT 2

ID	AAB50003	standard; protein; 890 AA.
----	----------	----------------------------

AC AAB50003;

DT 19-MAR-2001 (first entry)

Human METH2

Human, MET2; metalloprotease; thrombospondin; angiogenesis inhibition.  
 KW cancer therapy; benign tumour; ocular angiogenic disease;  
 KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;  
 KW vasculogenesis; granulation; hypertrophic scar; nonunion fracture;  
 KW scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;  
 KW coronary collateral; cerebral collateral; arteriovenous malformation;  
 KW ischaemic limb angiogenesis; Oler-Webber syndrome;  
 KW plaque neovascularisation; telangiectasia; haemophilic joint;  
 KW angiodiroma; fibromuscular dysplasia; wound granulation;  
 KW Crohn's disease; atherosclerosis; birth control.

OS Homo sapiens.

PN WO200071577-A1.

PD 30-NOV-2000.

PF 25-MAY-2000; 2000WO-US014462

PR 25-MAY-1999; 99US-00318208

PR	10-AUG-1999;	99US-0147823P
PR	10-AUG-1999;	99US-00373658

PR 22-DEC-1999; 99US-0171503F  
 PP 22-FEB-2000; 2000US-0183792P

XX  
PA (HTMA-) HITMAN GENOME SCI INC

PA (SMK ) SMITHLINE BEECHAM C  
PA (BETH-) BETH ISRAEL DEACONES

PA (KOE/) KUEH-ARISE U.  
PA (HAST/) HASTINGS G A.

PA (RUBE/) RUBEN S M.  
PA (JONA/) JONAK Z L.

PA (FORN/) FORNWALD J A.

XX  
XX  
XX

Formwald JA, Terrett JA; PI

DR WPI; 2001-025136/03.

XXXX

### PT angiogenesis in the treatment

XX

XXXXXX

thrombospandin). METH2 can b

CC disease, rheumatoid arthritis

CC fractures, scleroderma, tracheobronchitis, coronary colla-

CC malformations, ischaemic limb

aim 15; Fig 2; 768pp; English.

XX The present sequence is human METH2 (ME for metalloproteinase and TH for thrombospondin). METH2 can be used for inhibiting angiogenesis in an individual, and for treating cancer, benign tumours, an ocular angiogenic disease, rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis, vasculogenesis, granulomas, hypertrophic scars, fractures, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, coronary collaterals, Oler-Webber syndrome, plaque

CC neovascularisation, telangiectasia, haemophilic joints, angiodiroma,  
 CC fibromuscular dysplasia, wound granulation, Crohn's disease or  
 CC atherosclerosis. METH2 can also be used in birth control. METH2 can also  
 CC be used in diagnostic methods for the prognosis of cancer

XX Sequence 890 AA;

Query Match 100.0%; Score 890; DB 4; Length 890;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPPAPAPRMLPFLILLILLILLPLARGAPAPAGQASSELVPTRLPGSAGELALHLSA 60  
 DB 1 MPPAPAPRMLPFLILLILLILLPLARGAPAPAGQASSELVPTRLPGSAGELALHLSA 60  
 QY 61 FKGPTLRLAPDPSFLAPEFKIERLGSGRATGGERGLRGCFSGTVNGEPESLAAYSLC 120  
 DB 61 FKGPTLRLAPDPSFLAPEFKIERLGSGRATGGERGLRGCFSGTVNGEPESLAAYSLC 120  
 QY 121 RGLSGSFLDGEFTIQPQAGGSLAQPHRLQWGPAGARPLPRGPEWEVETGEGQGER 180  
 DB 121 RGLSGSFLDGEFTIQPQAGGSLAQPHRLQWGPAGARPLPRGPEWEVETGEGQGER 180  
 QY 121 RGLSGSFLDGEFTIQPQAGGSLAQPHRLQWGPAGARPLPRGPEWEVETGEGQGER 180  
 DB 121 RGLSGSFLDGEFTIQPQAGGSLAQPHRLQWGPAGARPLPRGPEWEVETGEGQGER 180  
 QY 181 GDHODESEESQEEAEAGASPEPPPLGATSRTRFVSEARFVETLLVADASAAAYGADL 240  
 DB 181 GDHODESEESQEEAEAGASPEPPPLGATSRTRFVSEARFVETLLVADASAAAYGADL 240  
 QY 241 QNHILLTMSVAARIYHPSIKSINLMVYKVLVEDEKMPVSDNGGLTLNFCNQMR 300  
 DB 241 QNHILLTMSVAARIYHPSIKSINLMVYKVLVEDEKMPVSDNGGLTLNFCNQMR 300  
 QY 301 FNGPSDRPHRYDTALLTRQNFQCGEGLCDTLGVADIGTICDPNKSCEVIDEGLOAH 360  
 DB 301 FNGPSDRPHRYDTALLTRQNFQCGEGLCDTLGVADIGTICDPNKSCEVIDEGLOAH 360  
 QY 361 TLAEHLGHTLSMHDSPKCTRLFGPMGKHWAAPFVLNQTLPSPCSAMYTELLDG 420  
 DB 361 TLAEHLGHTLSMHDSPKCTRLFGPMGKHWAAPFVLNQTLPSPCSAMYTELLDG 420  
 QY 361 TLAEHLGHTLSMHDSPKCTRLFGPMGKHWAAPFVLNQTLPSPCSAMYTELLDG 420  
 DB 361 TLAEHLGHTLSMHDSPKCTRLFGPMGKHWAAPFVLNQTLPSPCSAMYTELLDG 420  
 QY 421 GHGDCULDAAGALPLPTGLFGHMAIYOLDQCRQIFGDFRHCPTMSADQVCAIWCMT 480  
 DB 421 GHGDCULDAAGALPLPTGLFGHMAIYOLDQCRQIFGDFRHCPTMSADQVCAIWCMT 480  
 QY 481 DGAEPLCHTNGSLPMWADGTPCGPHLCSGSLCPBEVERKPVVDGMAHPMGEGS 540  
 DB 481 DGAEPLCHTNGSLPMWADGTPCGPHLCSGSLCPBEVERKPVVDGMAHPMGEGS 540  
 QY 541 RTGGGVQFSHRECKDEPQNGRYCLGRRAKYQSGHTEBCPPDGKSPBQCEKTNAYN 600  
 DB 541 RTGGGVQFSHRECKDEPQNGRYCLGRRAKYQSGHTEBCPPDGKSPBQCEKTNAYN 600  
 QY 601 YTDMDGNLLQWPKYAGVSPDRCKLFCRAAGSEFVFPAKYIDGLTQCFPTLACVAG 660  
 DB 601 YTDMDGNLLQWPKYAGVSPDRCKLFCRAAGSEFVFPAKYIDGLTQCFPTLACVAG 660  
 QY 661 QCVNAGCDHVVDSPKDKGCGVGGKSGSKRYSGSLTPNYGNDIVITIPAGATNIDYK 720  
 DB 661 QCVNAGCDHVVDSPKDKGCGVGGKSGSKRYSGSLTPNYGNDIVITIPAGATNIDYK 720  
 QY 721 QRSHPGVQNDGNLALKTADQYLLNGNLAIASIEODILVKGITLLKSGSIATLERIQSF 780  
 DB 721 QRSHPGVQNDGNLALKTADQYLLNGNLAIASIEODILVKGITLLKSGSIATLERIQSF 780  
 QY 781 RPLREPLTVOLLTPGVFPKPYKTYTFVPRVDVFSQSSKERATTNIIQPLHAQWVLG 840  
 DB 781 RPLREPLTVOLLTPGVFPKPYKTYTFVPRVDVFSQSSKERATTNIIQPLHAQWVLG 840  
 QY 841 DMSGCSSTCGAGWQRTVECRDPGQASATCNKALKEADAFCSQLCPL 890  
 DB 841 DMSGCSSTCGAGWQRTVECRDPGQASATCNKALKEADAFCSQLCPL 890

RESULT 3

ABP96306  
 ID ABP96306 standard; protein; 890 AA.  
 XX  
 AC ABP96306;  
 XX  
 DT 20-MAY-2003 (first entry)  
 XX  
 DE Human ADAMTS8 protein.  
 XX  
 KM Humanised baculovirus; cytosolic; gene therapy; baculovirus; cancer;  
 KM prostate cancer; chromosome 11.

OS Homo sapiens.  
 XX  
 PN W02003016540-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 15-AUG-2002; 2002WO-GB003791.  
 XX  
 PR 15-AUG-2001; 2001GB-00019852.  
 PA (UYO-) UNIV YORK.  
 XX  
 PI Maitland N;  
 XX  
 DR WPI; 2003-268336/26.  
 XX  
 PT New baculovirus having a modified genome encoding a therapeutic agent,  
 PT useful in the manufacture of a medicament for the treatment of cancer,  
 PT particularly prostate cancer.

PS Claim 24; Page; 34pp; English.

CC The present invention describes a humanised baculovirus (I) which  
 CC comprises a modified baculovirus genome having a nucleic acid molecule  
 CC encoding a therapeutic agent and a polypeptide which functions to target  
 CC the baculovirus to at least one cell type. Also described is a  
 CC pharmaceutical composition comprising (I). (I) has cytostatic activity,  
 CC and can be used in gene therapy. The baculovirus is useful in the  
 CC manufacture of a medicament for the treatment of cancer, particularly  
 CC prostate cancer. The present sequence represents the human disintegrin-  
 CC like and metalloprotease (repolysin type) with thrombospondin type 1  
 CC motif, 8 (ADMTS8) protein, which is specified in the exemplification of  
 CC the present invention as angiotensin. N.B. The present sequence is not  
 CC given in the specification but is referred to in Claim 24 as Genbank  
 CC accession number NM\_007037

XX Sequence 890 AA;

Query Match 75.8%; Score 675; DB 6; Length 890;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 875; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 LLLLLLLLLPLARGAPAPAGQASSELVPTRLPGSAGELALHLSAFKGFVRLAPDD 73  
 DB 14 LLLLLLLLLPLARGAPAPAGQASSELVPTRLPGSAGELALHLSAFKGFVRLAPDD 73  
 QY 14 LLLLLLLLLPLARGAPAPAGQASSELVPTRLPGSAGELALHLSAFKGFVRLAPDD 73  
 DB 14 LLLLLLLLLPLARGAPAPAGQASSELVPTRLPGSAGELALHLSAFKGFVRLAPDD 73  
 QY 74 SFLAEFKIERLGSGRATGGERGLRGCFSGTVNGEPESLAAYSLCRLGSGSFLDGE 133  
 DB 74 SFLAEFKIERLGSGRATGGERGLRGCFSGTVNGEPESLAAYSLCRLGSGSFLDGE 133  
 QY 134 FTIQPQAGGSLAQPHRLQWGPAGARPLPRGPEWEVETGEGQGERDHOEDSEESQ 193  
 DB 134 FTIQPQAGGSLAQPHRLQWGPAGARPLPRGPEWEVETGEGQGERDHOEDSEESQ 193  
 QY 194 EBAEGASEPPPLGATSRTRFVSEARFVETLLVADASAAAYGADLQNHILLTMSVAAR 253  
 DB 194 EBAEGASEPPPLGATSRTRFVSEARFVETLLVADASAAAYGADLQNHILLTMSVAAR 253  
 QY 254 IYKHPISIKSINLMVYKVLVEDEKMPVSDNGGLTLNFCNQMRDRPHRYD 313  
 DB 254 IYKHPISIKSINLMVYKVLVEDEKMPVSDNGGLTLNFCNQMRDRPHRYD 313

QY 314 TAILTRONFCGQEGICDTLGVADIGTICDPNKSCSVIEDEGLQAAMHTLAHELGHVLSMP 373  
 XX  
 CC 314 TAILTRONFCGQEGICDTLGVADIGTICDPNKSCSVIEDEGLQAAMHTLAHELGHVLSMP 373  
 QY 374 HDSPKCTRLFGPMGKHVMAPLFVHLNQTLPMSPCSAMTLTELDGSHDCLLDAPGAA 433  
 Db 374 HDSPKCTRLFGPMGKHVMAPLFVHLNQTLPMSPCSAMTLTELDGSHDCLLDAPGAA 433  
 QY 434 LPLPTGLPGRMALYQLDQOCROI FGPDRHCNPTSAQVCAQIMCHTDGAEPLCHTKNGS 493  
 Db 434 LPLPTGLPGRMALYQLDQOCROI FGPDRHCNPTSAQVCAQIMCHTDGAEPLCHTKNGS 493  
 QY 494 LPMADGTPCGPGLHLCSEGSCLPEEVEVERPKPVVDGMAFWGPGWGCSTRTGGGVQFSHRE 553  
 Db 494 LPMADGTPCGPGLHLCSEGSCLPEEVEVERPKPVVDGMAFWGPGWGCSTRTGGGVQFSHRE 553  
 QY 554 CKDEPONGGRYCLGRABKYOSCHTEBCPPDGKSFRRQOCCKYNAVYITMDGNLLQWPY 613  
 Db 554 CKDEPONGGRYCLGRABKYOSCHTEBCPPDGKSFRRQOCCKYNAVYITMDGNLLQWPY 613  
 QY 614 KYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLCGPETLALCVRGQCVKAGCDHVVD 673  
 Db 614 KYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLCGPETLALCVRGQCVKAGCDHVVD 673  
 QY 674 PRKLDKCGVCGGKNSCRKVSGLTPTNYGVNDIVTIPAGATNIDYKORSHPGVQNDGNY 723  
 Db 674 PRKLDKCGVCGGKNSCRKVSGLTPTNYGVNDIVTIPAGATNIDYKORSHPGVQNDGNY 723  
 QY 734 LALKTADQYLLNGNLAI SAIEDIIVKGTILKYSIATLERLOSFRPLPEPLTVQLLT 793  
 Db 734 LALKTADQYLLNGNLAI SAIEDIIVKGTILKYSIATLERLOSFRPLPEPLTVQLLT 793  
 QY 794 VPGEVPPPKYKTYTFVFNVDVFSMOSSEKERTNTIIQPLHQAOWVLGDMSECSSTCGAGW 853  
 Db 794 VPGEVPPPKYKTYTFVFNVDVFSMOSSEKERTNTIIQPLHQAOWVLGDMSECSSTCGAGW 853  
 QY 854 QRRTECRDPSGQASATCNKALKEPADKPCESQLCPL 890  
 Db 854 QRRTECRDPSGQASATCNKALKEPADKPCESQLCPL 890

RESULT 4  
 AAB74946  
 ID AAB74946 standard; protein; 889 AA.

AC AAB74946;  
 XX  
 DT 02-JUL-2001 (first entry)  
 DE Human ADAM type metal protease MDTs3 protein SEQ ID NO:20.  
 XX  
 KW Human; a disintegrin and metalloprotease type metal protease; MDTs1;  
 KW MDTs2; MDTs3; ADAM type metal protease; cytosolic; antiarthritic;  
 KW cancer; arthritis; arthrosis deformans.  
 XX  
 OS Homo sapiens.  
 PN JP2001008687-A.  
 PD 16-JAN-2001.  
 PF 25-JUN-1999; 99UP-00180973.  
 PR 25-JUN-1999; 99UP-00180973.  
 XX  
 PA (YAMA) YAMANOUCHI PHARM CO LTD.  
 XX  
 DR WPI, 2001-285362/30.  
 DR N-PSDB; AAF82166.  
 XX  
 PT New metal protease and metal protease gene, for use as a drug for  
 PT cancers, arthritis and arthrosis deformans.

XX  
 PS Claim 1; Page 25-27; 31pp; Japanese.  
 CC The present sequence represents a disintegrin and metalloprotease (ADAM)  
 CC type metal protease designated MDTs3, isolated from human. MDTs proteins  
 CC have cytosolic and antiarthritic activities. They can be used as a drug  
 CC for cancers, arthritis and arthrosis deformans  
 XX  
 SQ Sequence 889 AA;

Query Match 72.9%; Score 649; DB 4; Length 889;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 849; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 LILLILLILLIPARAPAPAAAGQASLIVPTRLPDGSAGELALHSAGKGFVRLAPDD 73  
 Db 13 LILLILLILLIPARAPAPAAAGQASLIVPTRLPDGSAGELALHSAGKGFVRLAPDD 72  
 QY 74 SFLADEFKIERLIGSGRATGERGLRGCFSGTVNGEPESLAAYSLCRGLSGSFLDGE 133  
 Db 73 SFLADEFKIERLIGSGRATGERGLRGCFSGTVNGEPESLAAYSLCRGLSGSFLDGE 132  
 QY 134 FTIQOGAGGSLAQPHRLQWGPAGARLPFGPEVETGSGORERDHDSEESOE 193  
 Db 133 FTIQOGAGGSLAQPHRLQWGPAGARLPFGPEVETGSGORERDHDSEESOE 192  
 QY 194 EEAEGSEPPPLGATSRKTFVSEARFETLLVADASMAAFYGADLQNHILTVMSVAR 253  
 Db 193 EEAEGSEPPPLGATSRKTFVSEARFETLLVADASMAAFYGADLQNHILTVMSVAR 252  
 QY 254 IYKHSIKNSINLMVYKLYVEDEKWPEVSDNGELTRNCNMORRNOGSDRPEHYD 313  
 Db 253 IYKHSIKNSINLMVYKLYVEDEKWPEVSDNGELTRNCNMORRNOGSDRPEHYD 312  
 QY 314 TAILTRONFCGQEGICDTLGVADIGTICDPNKSCSVIEDEGLQAAMHTLAHELGHVLSMP 373  
 Db 313 TAILTRONFCGQEGICDTLGVADIGTICDPNKSCSVIEDEGLQAAMHTLAHELGHVLSMP 372  
 QY 374 HDSPKCTRLFGPMGKHVMAPLFVHLNQTLPMSPCSAMTLTELDGSHDCLLDAPGAA 433  
 Db 373 HDSPKCTRLFGPMGKHVMAPLFVHLNQTLPMSPCSAMTLTELDGSHDCLLDAPGAA 432  
 QY 434 LPLPTGLPGRMALYQLDQOCROI FGPDRHCNPTSAQVCAQIMCHTDGAEPLCHTKNGS 493  
 Db 433 LPLPTGLPGRMALYQLDQOCROI FGPDRHCNPTSAQVCAQIMCHTDGAEPLCHTKNGS 492  
 QY 494 LPMADGTPCGPGLHLCSEGSCLPEEVEVERPKPVVDGMAFWGPGWGCSTRTGGGVQFSHRE 553  
 Db 493 LPMADGTPCGPGLHLCSEGSCLPEEVEVERPKPVVDGMAFWGPGWGCSTRTGGGVQFSHRE 552  
 QY 554 CKDEPONGGRYCLGRABKYOSCHTEBCPPDGKSFRRQOCCKYNAVYITMDGNLLQWPY 613  
 Db 553 CKDEPONGGRYCLGRABKYOSCHTEBCPPDGKSFRRQOCCKYNAVYITMDGNLLQWPY 612  
 QY 614 KYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLCGPETLALCVRGQCVKAGCDHVVD 673  
 Db 613 KYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLCGPETLALCVRGQCVKAGCDHVVD 672  
 QY 674 PRKLDKCGVCGGKNSCRKVSGLTPTNYGVNDIVTIPAGATNIDYKORSHPGVQNDGNY 723  
 Db 673 PRKLDKCGVCGGKNSCRKVSGLTPTNYGVNDIVTIPAGATNIDYKORSHPGVQNDGNY 732  
 QY 734 LALKTADQYLLNGNLAI SAIEDIIVKGTILKYSIATLERLOSFRPLPEPLTVQLLT 793  
 Db 733 LALKTADQYLLNGNLAI SAIEDIIVKGTILKYSIATLERLOSFRPLPEPLTVQLLT 792  
 QY 794 VPGEVPPPKYKTYTFVFNVDVFSMOSSEKERTNTIIQPLHQAOWVLGDMSECSSTCGAGW 853  
 Db 793 VPGEVPPPKYKTYTFVFNVDVFSMOSSEKERTNTIIQPLHQAOWVLGDMSECSSTCGAGW 852  
 QY 854 QRRTECRDPS 864  
 Db 853 QRRTECRDPS 863



RESULT 5  
 AAB21251  
 ID AAB21251 standard; protein; 660 AA.  
 XX  
 AC AAB21251;  
 XX  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Human metalloproteinase ADAMTS-2.  
 XX  
 KW Human; ADAMTS2; metalloproteinase; ADAM;  
 KW a disintegrin and metalloproteinase domain; thrombospondin domain;  
 KW vaccine; nootropic; neuroprotective; antiparkinsonian; cerebroprotective;  
 KW cytoskeletal; antiarthritic; immunosuppressive; Alzheimer's disease;  
 KW Parkinson's disease; stroke; cancer; arthritis; autoimmune disease;  
 KW brain tumour; brain injury.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200053774-A2.  
 XX  
 PD 14-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000MO-US006237.  
 XX  
 PR 08-MAR-1999; 99US-00264585.  
 XX  
 PA (NEUR-) NEUROCRINE BIOSCIENCES INC.  
 XX  
 PI Keiner GS, Clark M, Maki RA;  
 XX  
 DR WPI; 2000-594326/56.  
 DR N-PSDB; AAA95821.  
 XX  
 PT Polymorphic encoding novel members of a disintegrin, metalloproteinase  
 PT and thrombospondin domain protein family used to prevent and treat  
 PT Alzheimer's disease, cancer and autoimmune diseases.  
 XX  
 PS Claim 12; Fig 2; 129pp; English.  
 XX  
 CC The present sequence is human metalloproteinase ADAMTS-2. The ADAMTS  
 CC family of proteins is closely related to the ADAM (A Disintegrin and  
 CC Metalloproteinase Domain) family. Members of the ADAMTS family contain a  
 CC thrombospondin domain in addition to the disintegrin and  
 CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are  
 CC useful for the manufacture of medicaments for treating conditions  
 CC associated with neuroinflammation and/or neurodegeneration, such as  
 CC Alzheimer's disease, Parkinson's disease and stroke. They are also useful  
 CC for treating conditions associated with cell proliferation, cell  
 CC migration, inflammation and/or angiogenesis, such as cancer, arthritis  
 CC and autoimmune diseases. They can be used to treat patients afflicted  
 CC with an invasive tumour, a brain tumour or brain injury  
 XX  
 SQ Sequence 680 AA;  
 Query Match 40.7%; Score 362; DB 3; Length 680;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 462; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 391 HMAAPLVPHINQTLPMSPCSAMYITELLIDGHHGDDLLDAPGALPLPTGLFGRMALYQLD 450  
 DB 181 HMAAPLVPHINQTLPMSPCSAMYITELLIDGHHGDDLLDAPGALPLPTGLFGRMALYQLD 240  
 QY 451 QOCROIFGDPFRHCPNPSADQVCAQLWCHTGAEPICHTKXGSLPMADGTPCGGHLTCE 510  
 DB 241 QOCROIFGDPFRHCPNPSADQVCAQLWCHTGAEPICHTKXGSLPMADGTPCGGHLTCE 300  
 QY 511 GSCLPSEEVERPKXVVDGMAFMGPKWGCSTTCGGGQVQFHSRECKDPEPONGRYCLGRR 570  
 DB 301 GSCLPSEEVERPKXVVDGMAFMGPKWGCSTTCGGGQVQFHSRECKDPEPONGRYCLGRR 360  
 QY 571 AKYOSCHTEECPPGPKSPREOQCEKXNAVNTDMDGNLQWPKYAGVSPDRCKLFCRA 630  
 DB 361 AKYOSCHTEECPPGPKSPREOQCEKXNAVNTDMDGNLQWPKYAGVSPDRCKLFCRA 420  
 QY 631 RGRSEFKYFEAKVIDGTLGPEETLAICVRGQCVAGCDHVDS 673  
 DB 421 RGRSEFKYFEAKVIDGTLGPEETLAICVRGQCVAGCDHVDS 463

RESULT 6  
 AAE09710  
 ID AAE09710 standard; protein; 364 AA.  
 XX  
 AC AAE09710;  
 XX  
 DT 22-NOV-2001 (first entry)  
 XX  
 DE Human gene 6 encoding novel protein HCE4D69, SEQ ID NO:57.  
 XX  
 KW Human; cytoskeletal; gene therapy; inflammatory disorder; neural disorder;  
 KW Parkinson's disease; Alzheimer's disease; immune system disorder; AIDS;  
 KW autoimmune disease; rheumatoid arthritis; muscular disorder; ischaemia;  
 KW reproductive disorder; Crohn's disease; pulmonary disorder; cancer;  
 KW myocardial infarction; glomerulonephritis; nephrotic syndrome; tumour;  
 KW haematopoietic disorder; rhinitis; asthma; diabetes; atherosclerosis;  
 KW anti-sense therapy; endocrine disorder; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155202-A1.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US001325.  
 XX  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.  
 PR 14-AUG-2000; 2000US-0225267P.  
 PR 14-AUG-2000; 2000US-0225268P.  
 PR 14-AUG-2000; 2000US-0225270P.

[illegible]

08-NOV-2000; 2000US-0246532P.  
08-NOV-2000; 2000US-0246603P.  
08-NOV-2000; 2000US-0246610P.  
08-NOV-2000; 2000US-0246611P.  
08-NOV-2000; 2000US-0246613P.  
17-NOV-2000; 2000US-0249207P.  
17-NOV-2000; 2000US-0249208P.  
17-NOV-2000; 2000US-0249209P.  
17-NOV-2000; 2000US-0249210P.  
17-NOV-2000; 2000US-0249211P.  
17-NOV-2000; 2000US-0249212P.  
17-NOV-2000; 2000US-0249213P.  
17-NOV-2000; 2000US-0249214P.  
17-NOV-2000; 2000US-0249215P.  
17-NOV-2000; 2000US-0249216P.  
17-NOV-2000; 2000US-0249217P.  
17-NOV-2000; 2000US-0249218P.  
17-NOV-2000; 2000US-0249244P.  
17-NOV-2000; 2000US-0249245P.  
17-NOV-2000; 2000US-0249264P.  
17-NOV-2000; 2000US-0249265P.  
17-NOV-2000; 2000US-0249297P.  
17-NOV-2000; 2000US-0249299P.  
17-NOV-2000; 2000US-0249300P.  
01-DEC-2000; 2000US-0250160P.  
01-DEC-2000; 2000US-0250391P.  
05-DEC-2000; 2000US-0251030P.  
05-DEC-2000; 2000US-0251988P.  
05-DEC-2000; 2000US-0256719P.  
06-DEC-2000; 2000US-0251479P.  
08-DEC-2000; 2000US-0251856P.  
08-DEC-2000; 2000US-0251868P.  
08-DEC-2000; 2000US-0251869P.  
08-DEC-2000; 2000US-0251899P.  
08-DEC-2000; 2000US-0251990P.  
11-DEC-2000; 2000US-0254097P.  
05-JAN-2001; 2001US-0255678P.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI: 2001-451925/48.  
N-PSDB; AAD16770.  
Isolated polypeptide for treating, preventing and/or prognosing medical disorders and also for testing and detection e.g. diagnosis and screening for agonists.  
Claim 11; SEQ ID NO 57; 469pP; English.  
AA016750-AA016775 represent cDNAs corresponding to novel human protein genes, and AAB09690-AAB09715 represent the proteins they encode. AA016777-ADAD16780 represent novel human genomic DNA fragments. The novel proteins and their DNAs are useful for diagnosing, treating, preventing and/or prognosing inflammatory disorders (bursitis or tendinitis); neural disorders (e.g. Parkinson's disease, Alzheimer's disease); immune system disorders; AIDS; autoimmune diseases (e.g., rheumatoid arthritis); muscular disorders; reproductive disorders; gastrointestinal disorders; (malabsorption syndrome, Crohn's disease); pulmonary disorders; cardiovascular disorders (myocardial infarction, ischaemia, arrhythmias); renal disorders (glomerulonephritis, nephrotic syndrome); cancerous disease and conditions (breast cancer); hyperproliferative disorders (leukemia, hyperplasia); tumours; foetal and developmental abnormalities; haematopoietic disorders; respiratory disorders (rhinitis, asthma); angiogenic disorders; diabetes; atherosclerosis; endocrine disorders; pregnancy-related disorders and infections. The novel protein DNA is useful in gene therapy and anti-sense therapy. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage

CC properties. The present sequence represents a novel human protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences

Query Match 29.6%; Score 263; DB 4; Length 364;  
Best Local Similarity 100.0%; Pred. No. 4.4e-252;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 530 WAPMGWGWGSCSRCTGGGVOFSHRECKDPFPONGKRYCLGRRAKYOSCHTEECPPPGKSPR 589  
DB 4 WAPMGWGWGSCSRCTGGGVOFSHRECKDPFPONGKRYCLGRRAKYOSCHTEECPPPGKSPR 63  
QY 590 EOCCEKXNANYNYDMGNLQWPKYAGVSPDRCLFCFRANGSEPFKFEAKVIDGTLIC 649  
DB 64 EOCCEKXNANYNYDMGNLQWPKYAGVSPDRCLFCFRANGSEPFKFEAKVIDGTLIC 123  
QY 650 GPETLACVKGQCVKAGCDHYVSPRLKCKYCGKGNKSGSLPTNYGYNDIVT 709  
DB 124 GPETLACVKGQCVKAGCDHYVSPRLKCKYCGKGNKSGSLPTNYGYNDIVT 183  
QY 710 IPAGATNIDVKQSHGPGVNDGNYLAKTADGQYLNGNLAIASIEODILVKTILKXSG 769  
DB 184 IPAGATNIDVKQSHGPGVNDGNYLAKTADGQYLNGNLAIASIEODILVKTILKXSG 243  
QY 770 SIATLERLOSFRPLPEPLTVQLL 792  
DB 244 SIATLERLOSFRPLPEPLTVQLL 266  
RESULT 7  
ADCC2088  
ID ADCC2088 standard; protein; 364 AA.  
XX  
AC ADCC2088;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human protein from secreted protein gene 6 #2.  
XX  
KW Secreted protein; cytosolic; antibacterial; virucide; neuroprotective;  
KW gynascological; gastrointestinal; Gen; cardiac; cardiovascular; Gen;  
KW nephrotoxic; anti-inflammatory; muscular; Gen; respiratory; Gen;  
KW immunosuppressive; cerebroprotective; vasotropic; nocotropic;  
KW anti-allergic; cancer; bacterial infection; viral infection;  
KW neural disorder; immune system disorder; blood disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW inflammatory disorder; proliferative disorder; Human.  
XX  
OS Homo sapiens.  
XX  
PN US2003082681-A1.  
XX  
PD 01-MAY-2003.  
XX  
PF 07-MAR-2002; 2002US-00091391.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 07-JUN-2000; 2000US-0209515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218230P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225269P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225478P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232387P.  
PR 14-SEP-2000; 2000US-0232388P.  
PR 14-SEP-2000; 2000US-0232389P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236357P.  
PR 29-SEP-2000; 2000US-0236357P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 08-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251865P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 PR 17-JAN-2001; 2001US-00764903.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Ruben SM, Barash SC;  
 PI  
 XX  
 XX MPI; 2003-786903/74.  
 DR N-PSDB; ADOC22062.  
 XX  
 PT New nucleic acid molecules and polypeptides for diagnosing, preventing or  
 PT treating disorders associated with aberrant expression of the  
 PT polypeptide, e.g. neural or cardiovascular disorders, and in chromosome  
 PT identification.  
 XX  
 PS Claim 11; SEQ ID NO 57; 242bp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid molecule (cDNA)  
 CC encoding a human secreted protein, representing one of 15 novel genes.  
 CC Also included are recombinant vectors, host cells (expressing the  
 CC protein), the secreted proteins (including their fragments, epitopes and  
 CC homologues), an isolated antibody that binds specifically to the protein,  
 CC diagnosing a pathological condition or susceptibility to a pathological  
 CC condition (comprising determining the presence or absence of a mutation  
 CC in the nucleic acid and diagnosing a condition based on the presence or  
 CC absence of the mutation), diagnosing a pathological condition or

CC susceptibility to a pathological condition (comprising determining the  
 CC presence or amount of expression of the protein in a biological sample  
 CC and diagnosing a condition based on the presence or amount of expression  
 CC of the protein), preventing, treating or ameliorating a medical condition  
 CC by administering the nucleic acid or protein to a mammalian subject,  
 CC identifying a binding partner to the protein, the gene corresponding to  
 CC the cDNA sequence, and identifying an activity in a biological assay  
 CC (comprising expressing the nucleic acid in a cell, isolating the  
 CC supernatant, detecting an activity in a biological assay and identifying  
 CC the protein in the supernatant having the activity). The nucleic acids  
 CC and proteins display the following activities: Cytostatic, antibacterial,  
 CC Viricidal, Neuroprotective, Gynaecological, Gastrointestinal, Gen,  
 CC Cardiac, Cardiovascular-Gen, Nephrotoxic, Antiinflammatory, Muscular-  
 CC  
 Query Match 29.6%; Score 263; DB 7; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-252;  
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 530 WAPMGPMGECSPRTGGVGFHRECKDEPONGRCYICGRPAKIQSCITECPDQKSF 585  
 DB 4 WAPMGPMGECSPRTGGVGFHRECKDEPONGRCYICGRPAKIQSCITECPDQKSF 63  
 QY 590 ECGCEKNAVNTDMDGNTLWVPEKAVSPRDKLFCRARGSEEFVFAKVIDGLIC 649  
 DB 64 ECGCEKNAVNTDMDGNTLWVPEKAVSPRDKLFCRARGSEEFVFAKVIDGLIC 123  
 QY 650 GPETLAICVRGQCVAGCDHVVDSEPRKLDKCGVCGKNSCRKVSGLTPTNYGNDIVT 705  
 DB 124 GPETLAICVRGQCVAGCDHVVDSEPRKLDKCGVCGKNSCRKVSGLTPTNYGNDIVT 183  
 QY 710 IPAGANTIDVQRSHPGVONGNYALATKAPAGVILNGNLAISAIEDIILVGTILKTS 769  
 DB 184 IPAGANTIDVQRSHPGVONGNYALATKAPAGVILNGNLAISAIEDIILVGTILKTS 243  
 QY 770 SIATLERLOSFRPLPEPLTVQL 792  
 DB 244 SIATLERLOSFRPLPEPLTVQL 266  
 RESULT 8  
 ID AAB72285 standard; protein; 245 AA.  
 XX AAB72285;  
 AC AAB72285;  
 XX  
 XX 14-MAY-2001 (first entry)  
 XX  
 XX Human ADAMTS-8 amino acid sequence.  
 DE  
 XX ADAMTS-N, disintegrin; metalloprotease; thrombospondin type I motif;  
 XX tumour cachexia; inflammation; dermatosparaxis; EDS-VIC; angiogenesis;  
 KW Ehlers-Danlos syndrome type VIIC; articular cartilage erosion; human;  
 KW metastasis; embryogenesis; egg implantation; ADAMTS-8.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX MO200111074-A2.  
 XX  
 XX 15-FEB-2001.  
 PD  
 XX 03-AUG-2000; 2000MO-US021223.  
 PF  
 XX  
 XX 06-AUG-1999; 99US-00369364.  
 PR  
 XX (CLEV-) CLEVELAND CLINIC FOUND.  
 PA (APTE/) APTE S S.  
 PA (HURS/) HURSKATINEN T L.  
 PA (HIRO/) HIROHARA S.  
 XX  
 XX Apte SS, Hurskainen TL, Hirohata S;  
 PI  
 XX MPI; 2001-159978/16.  
 DR N-PSDB; AAF63442.

XX Murine and human 'A disintegrin-like And Metalloprotease domain with  
PT Thrombospondin type I motifs' proteins and the nucleic acids encoding  
PT them, useful for treating e.g. tumors, inflammation and arthritis.  
XX  
XX Claim 1; Fig 6; 181pp; English.  
XX  
XX This invention relates to murine and human ADAMTS-N (A disintegrin-like  
CC and metalloprotease domain with thrombospondin type I motifs) proteins,  
CC designated ADAMTS-5, 6, 7, 8, 9, 10 and R1. Also included in the  
CC invention are cDNA sequences encoding the proteins, and antibodies  
CC specific for the proteins. The nucleic acid sequences and proteins may be  
CC used in the prevention, diagnosis and treatment of diseases associated  
CC with inappropriate ADAMTS-N expression. Disorders that may be treated  
CC using the nucleic acids, proteins and antibodies include, for example  
CC tumour cachexia, inflammation, dermatosparaxis in cattle or Ehlers-Danlos  
CC syndrome type VIIC (EDS-VIIC) in humans, erosion of articular cartilage  
CC in arthritic (both inflammatory and non-inflammatory) disease,  
CC angiogenesis, tumour growth and metastases, and they may also be used for  
CC controlling embryogenesis and implantation of fertilised eggs. The  
CC present sequence represents human ADAMTS-8  
XX  
SQ Sequence 245 AA:  
Query Match 24.4%; Score 217; DB 4; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.6e-206;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 196 AEGASEPPPLGATSRTRKRVSEARFVETLLVADASMAAFYAGDLQNHITLMSVARIY 255  
DB 2 AEGASEPPPLGATSRTRKRVSEARFVETLLVADASMAAFYAGDLQNHITLMSVARIY 61  
XX  
XX 256 KHPISKINSLMVKYKILVDEKMGKPVSDNGGILRNPNMGRFRNPSDRHPEHYDA 315  
DB 62 KHPISKINSLMVKYKILVDEKMGKPVSDNGGILRNPNMGRFRNPSDRHPEHYDA 121  
XX  
XX 316 ILLTRONFCGEGCLDTLGVADIGTICDPNKSQVIEDEGLQAAHTLAHGLVLSMPHD 375  
DB 122 ILLTRONFCGEGCLDTLGVADIGTICDPNKSQVIEDEGLQAAHTLAHGLVLSMPHD 181  
XX  
XX 376 DSKPCTRLFGPMGKHVWAPLFVHLNQLTLPWSPCSAM 412  
DB 182 DSKPCTRLFGPMGKHVWAPLFVHLNQLTLPWSPCSAM 218  
XX  
XX  
XX RESULT 9  
XX AAB72284  
XX ID AAB72284 standard; protein; 905 AA.  
XX  
XX AAB72284;  
XX  
XX 14-MAY-2001 (first entry)  
XX  
XX Murine ADAMTS-8 amino acid sequence.  
XX  
XX ADAMTS-N; disintegrin; metalloprotease; thrombospondin type I motif;  
XX tumour cachexia; inflammation; dermatosparaxis; EDS-VIIC; angiogenesis;  
XX Ehlers-Danlos syndrome type VIIC; articular cartilage erosion; mouse;  
XX metastasis; embryogenesis; egg implantation; ADAMTS-8.  
XX  
XX Mus musculus.  
XX  
XX WO20011074-A2.  
XX  
XX 15-FEB-2001.  
XX  
XX 03-AUG-2000; 2000WO-US021223.  
XX  
XX 06-AUG-1999; 99US-00369364.  
XX  
XX (CLEV-) CLEVELAND CLINIC FOUND.  
XX (APTE-) APTE S S.  
XX (HURS/) HURSKAINEN T L.

PA (HIRO/) HIROHATA S.  
XX  
XX Ape SS, Hurskainen TL, Hirohata S;  
PI  
XX WPI; 2001-159978/16.  
XX  
XX DR N-PSDB; AAF63441.  
XX  
XX Murine and human 'A Disintegrin-like And Metalloprotease domain with  
PT Thrombospondin type I motifs' proteins and the nucleic acids encoding  
PT them, useful for treating e.g. tumors, inflammation and arthritis.  
XX  
XX Claim 1; Fig 5; 181pp; English.  
XX  
XX This invention relates to murine and human ADAMTS-N (A disintegrin-like  
CC and metalloprotease domain with thrombospondin type I motifs) proteins,  
CC designated ADAMTS-5, 6, 7, 8, 9, 10 and R1. Also included in the  
CC invention are cDNA sequences encoding the proteins, and antibodies  
CC specific for the proteins. The nucleic acid sequences and proteins may be  
CC used in the prevention, diagnosis and treatment of diseases associated  
CC with inappropriate ADAMTS-N expression. Disorders that may be treated  
CC using the nucleic acids, proteins and antibodies include, for example  
CC tumour cachexia, inflammation, dermatosparaxis in cattle or Ehlers-Danlos  
CC syndrome type VIIC (EDS-VIIC) in humans, erosion of articular cartilage  
CC in arthritic (both inflammatory and non-inflammatory) disease,  
CC angiogenesis, tumour growth and metastases, and they may also be used for  
CC controlling embryogenesis and implantation of fertilised eggs. The  
CC present sequence represents murine ADAMTS-8  
XX  
SQ Sequence 905 AA:  
Query Match 3.9%; Score 35; DB 4; Length 905;  
Best Local Similarity 100.0%; Pred. No. 3.8e-25;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 617 GVSPPDRCKLFCRARGSEPFVFAKVTGTLCP 651  
DB 632 GVSPPDRCKLFCRARGSEPFVFAKVTGTLCP 666  
XX  
XX  
XX RESULT 10  
XX AA04145  
XX ID AA04145 standard; protein; 481 AA.  
XX  
XX AA04145;  
XX  
XX 15-JUN-1999 (first entry)  
XX  
XX Rat Tango-76 protein.  
XX  
XX Rat; Tango-71; Tango-73; Tango-74; Tango-76; Tango-83; diagnosis;  
XX detection.  
XX  
XX Rattus sp.  
XX  
XX WO9907850-A1.  
XX  
XX 18-FEB-1999.  
XX  
XX 06-AUG-1998; 98WO-US016502.  
XX  
XX 06-AUG-1997; 97US-0054966P.  
XX  
XX 05-SEP-1997; 97US-0058108P.  
XX  
XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
XX  
XX Holtzman DA, Goodearl ADJ;  
XX  
XX WPI; 1999-167426/14.  
XX  
XX N-PSDB; AAX19958.  
XX  
XX New TANGO polypeptides and nucleic acids encoding them - useful as  
PT diagnostic agents and for treating disorders caused by aberrant  
PT expression of TANGO.

XX Claim 8; Fig 6; 84pp; English.

XX The present sequence represents rat Tango-71. Tango polypeptides are

CC useful for identifying compounds which bind the polypeptide via direct

CC binding, competition binding assays or Tango-71, -73, -74, 76 or -83-

CC mediated signal transduction. Tango polypeptides are also useful for

CC identifying modulating compounds by determining effect on Tango activity.

CC Tango polypeptides and nucleic acids are useful for diagnosing diseases

CC related to aberrant expression of Tango, and Tango polypeptides are

CC useful for raising antibodies which can be used in diagnostic assays for

CC detection of Tango, and also for generating anti-idiotypic antibodies for

CC prevention and protection

XX Sequence 481 AA:

Query Match 3.4%; Score 30; DB 2; Length 481;

Best Local Similarity 100.0%; Pred. No. 2e-20; Mismatches 0; Indels 0; Gaps 0;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 702 YGNDIVTTPAGATNIDVKQRSHPGVQNDG 731

DB 293 YGNDIVTTPAGATNIDVKQRSHPGVQNDG 322

RESULT 11

AAB21262

ID AAB21262 standard; protein; 98 AA.

XX AAB21262;

XX 23-FEB-2001 (first entry)

DE Rat metalloproteinase ADAMTS-2.

XX Rat; ADAMTS-2; metalloproteinase; ADM;

KW a disintegrin and metalloproteinase domain; thrombospondin domain;

KW vaccine; nootropic; neuroprotective; antiparkinsonian; cerebroprotective;

KW cytosolic; anticholinergic; immunosuppressive; Alzheimer's disease;

KW Parkinson's disease; stroke; cancer; arthritis; autoimmune disease;

KW brain tumour; brain injury.

XX Rattus norvegicus.

OS WO200053774-A2.

PN 14-SEP-2000.

XX 08-MAR-2000; 2000WO-US006237.

PF 08-MAR-1999; 99US-00264585.

XX (NEUR-) NEUROCRINE BIOSCIENCES INC.

PA Kellner GS, Clark M, Maki RA;

PI WPI; 2000-594326/56.

DR N-PSDB; AAA95832.

XX Polynucleotide encoding novel members of a disintegrin, metalloproteinase

PT and thrombospondin domain protein family used to prevent and treat

PT Alzheimer's disease, cancer and autoimmune diseases.

XX Claim 12; Fig 25; 129pp; English.

XX The present sequence is rat metalloproteinase ADAMTS-2. The ADAMTS family

CC of proteins is closely related to the ADM (a disintegrin and

CC metalloproteinase domain) family. Members of the ADAMTS family contain a

CC thrombospondin domain in addition to the disintegrin and

CC metalloproteinase domain found in the ADMs. ADAMTS polypeptides are

CC useful for the manufacture of medicaments for treating conditions

CC associated with neuroinflammation and/or neurodegeneration, such as

CC Alzheimer's disease, Parkinson's disease and stroke. They are also useful

CC for treating conditions associated with cell proliferation, cell

CC migration, inflammation and/or angiogenesis, such as cancer, arthritis

CC and autoimmune diseases. They can be used to treat patients afflicted

CC with an invasive tumour, a brain tumour or brain injury

XX Sequence 98 AA:

Query Match 2.7%; Score 24; DB 3; Length 98;

Best Local Similarity 100.0%; Pred. No. 4.4e-15; Mismatches 0; Indels 0; Gaps 0;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 617 GVSPPDRCKLFCRARGSEFKYFE 640

DB 32 GVSPPDRCKLFCRARGSEFKYFE 55

RESULT 12

AAY53898

ID AAY53898 standard; protein; 2150 AA.

XX AAY53898;

XX 13-MAR-2000 (first entry)

DE Amino acid sequence of a GON-1 protein of *Caenorhabditis elegans*.

XX GON-1; metalloproteinase; cell migration; modulator;

KW metalloproteinase domain; thrombospondin domain; abnormal cell migration;

KW organ shaping; sterility; cancer metastasis.

XX *Caenorhabditis elegans*.

OS Key Location/Qualifiers

FH Domain 269..456

FT /note="metalloprotease domain"

XX WO9961656-A2.

PN 02-DEC-1999.

XX 28-MAY-1999; 99WO-US011918.

PF 29-MAY-1998; 98US-0087170P.

PR 13-APR-1999; 99US-0129023P.

XX (WISC) WISCONSIN ALUMNI RES FOUND.

PA Kimble JE, Bleiloch RH;

PI WPI; 2000-072633/06.

DR N-PSDB; AAZ36862.

XX Identifying modulators of proteins containing metalloprotease and

PT thrombospondin domains, potentially useful for controlling cell migration

PT and organ shaping.

XX Disclosure; Page 51-60; 60pp; English.

XX The present sequence represents GON-1 protein of *Caenorhabditis elegans*.

CC GON-1 is a secreted metalloproteinase that lacks a transmembrane domain

CC and possesses a predicted metalloprotease domain between amino acids 269-

CC 456. In *C. elegans* hermaphrodites, GON-1 is required for migration of two

CC distal tip cells to produce elongated tubes, whereas in males, GON-1 is

CC required for migration of a single linker cell to produce a single

CC elongated tube. The protein is used in the method of the invention. The

CC specification describes a method for identifying a modulator of a protein

CC that contains a metalloproteinase domain and a thrombospondin domain. The

CC method comprises treating a target organism, having a developing gonadal

CC cell that is responsive to the protein, with a test compound, and

CC determining any change in migration or shape of the cell attributable to

CC the test compound. The compounds identified are potential therapeutic

CC modulators of abnormal cell migration and organ shaping, e.g. for

CC rendering animals (specifically nematodes) sterile and for inhibiting

CC Cancer metaetases  
 XX  
 SQ Sequence 2150 AA;

Query Match 1.5%; Score 13; DB 3; Length 2150;  
 Best Local Similarity 100.0%; Pred. No. 0.0061;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGAVQ 548  
 |||||  
 DB 597 WGECSRTCGGAVQ 609

RESULT 13  
 AAB90617  
 ID AAB90617 standard; protein; 2165 AA.  
 XX  
 AC AAB90617;  
 XX  
 DT 01-JUN-2001 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID NO: 155.  
 XX  
 XX Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
 KW antiinflammatory; anti-HIV; cytostatic; cardiac; vascular;  
 KW anti-angiogenic; ophthalmological; neuroprotectant; nootropic;  
 KW anticonvulsant; antialzheimers; antiparkinsonian; antitubercial;  
 KW vulnerary; vaccine; gene therapy; cancer; protein coordinate data;  
 XX infection.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WC200121658-A1.  
 PD  
 XX 29-MAR-2001.  
 PF  
 XX 22-SEP-2000; 2000MO-US026013.  
 XX  
 PR 24-SEP-1999; 99US-0155709P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCT INC.  
 XX  
 PI Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;  
 PI Lafleur DM, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;  
 PI Young PE, Wei P, Florence KA;  
 XX  
 DR WPI: 2001-235311/24.

Nucleic acids encoding 32 human secreted polypeptides, useful for preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy.

PS Disclosure; Page 830-836; 890pp; English.

XX The present sequence is provided in a specification relating to nucleic  
 CC acid molecules encoding 32 novel human secreted polypeptides. The nucleic  
 CC acid molecules and polypeptides may be used in the prevention, diagnosis  
 CC and treatment of diseases such as immune disorders (e.g. multiple  
 CC sclerosis, systemic lupus erythematosus and human immuno-deficiency virus  
 CC (HIV) infections), hyperproliferative disorders (e.g. cancers and  
 CC Gaucher's disease), cardiovascular diseases (e.g. Schmitz syndrome,  
 CC Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic  
 CC disorders (e.g. corneal graft neovascularization and diabetic  
 CC retinopathy), neurological disorders (e.g. Huntington's chorea,  
 CC Alzheimer's disease and Parkinson's disease), infectious diseases and/or  
 CC for promoting wound healing, regeneration and/or chemotaxis. The nucleic  
 CC acid molecules may be used to produce the secreted polypeptides. They may  
 CC also be used as DNA probes in diagnostic assays to detect and quantitate  
 CC the presence of similar nucleic acid sequences in samples. The  
 CC polypeptides may be used as antigens in the production of antibodies and  
 CC in assays to identify modulators of their expression and activity  
 SQ Sequence 2165 AA;

Query Match 1.5%; Score 13; DB 4; Length 2165;  
 Best Local Similarity 100.0%; Pred. No. 0.0061;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGAVQ 548  
 |||||  
 DB 612 WGECSRTCGGAVQ 624

RESULT 14  
 AAP81021  
 ID AAP81021 standard; protein; 18 AA.  
 XX  
 AC AAP81021;  
 XX  
 DT 15-NOV-1990 (first entry)  
 XX  
 DE Sequence of modified egg white lysozyme signal peptide.  
 XX  
 XX Modified egg white lysozyme signal sequence; enzyme.  
 KW  
 OS Homo sapiens.  
 XX  
 PN JP63233789-A.  
 PD  
 XX 29-SEP-1988.  
 PF  
 XX 23-MAR-1987; 87JP-00069764.  
 PR  
 XX 23-MAR-1987; 87JP-00069764.  
 XX  
 PA (AJIN ) TANPAKU KAGAKU KENK.  
 XX  
 DR WPI: 1988-318076/45.  
 DR N-PSDB; AAP81348.  
 XX  
 PT DNA sequence encoding signal peptide - having good protein secretion  
 PT activity.  
 XX  
 PS Claim 3; Page 523; 11pp; Japanese.  
 XX  
 CC AAs of natural egg white lysozyme other than residues 1, 2 and 16-18 are  
 CC all or mostly substituted with one kind of hydrophobic AA. Using the  
 CC modified signal peptide, foreign proteins, esp. human lysozyme, can be  
 CC secreted correctly and at a high level  
 XX  
 SQ Sequence 18 AA;

Query Match 1.3%; Score 12; DB 1; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 0.00082;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLLPLA 25  
 |||||  
 DB 4 LLLLLLLLLPLA 15

RESULT 15  
 AAB97107  
 ID AAB97107 standard; peptide; 19 AA.  
 XX  
 AC AAB97107;  
 XX  
 DT 02-AUG-2001 (first entry)  
 XX  
 DE Chlorella signal peptide.  
 XX  
 KW Chlorella; signal peptide; gene expression; protein production;  
 KW human growth hormone.  
 XX  
 OS Chlorella sp.  
 XX



Fri May 7 12:17:59 2004

us-09-989-687-4.rag

Page 12

PN JP2000354490-A.  
XX  
PD 26-DEC-2000.  
XX  
PF 15-JUN-1999; 99JP-00168271.  
XX  
PR 15-JUN-1999; 99JP-00168271.  
XX  
PA (TOYT ) TOYOTA JIDOSHA KK.  
XX  
DR WPI; 2001-275809/29.  
XX  
PT New signal peptides useful for the preparation of human growth hormone  
PT and transformed chlorella.  
XX  
PS Disclosure; Page 12; 15pp; Japanese.  
XX  
CC The present sequence is provided in a specification relating to signal  
CC peptides for expression and secretion of a protein in chlorella. The  
CC peptides are of the formula: Met-Ala-Asn-Lys-X<sub>1</sub>-(Leu)-n-X<sub>2</sub>-Ala-Ser-Gly.  
CC X<sub>1</sub> = Ser or Leu; n = an integer of 5-15; X<sub>2</sub> = Gly-Ser-Leu or Pro-Leu-  
CC Ala. The signal peptides are useful in the preparation of human growth  
CC hormone and transformed chlorella. Signal peptides, DNA encoding the  
CC peptides, gene expression cassettes, recombinant vectors containing the  
CC cassettes, and transformants having the vectors are provided  
XX  
SQ Sequence 19 AA;  
Query Match 1.3%; Score 12; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.00086;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 14 LLLLLLLLLPLA 25  
DB 5 LLLLLLLLLPLA 16

Search completed: May 7, 2004, 11:58:36  
Job time : 68 secs

Fri May 7 12:18:02 2004

us-09-989-687-4\_1.rapb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:51:11; Search time 54 Seconds  
(without Alignment)  
4574.711 Million cell updates/sec

Title: US-09-989-687-4  
Sequence: 1 MFPAAPAPWLPFLILLILL.....CNKALKPDAKPCESQCLPL 890

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1140673 segs, 27566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PC1\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PC1US\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4853	100.0	890	US-09-373-658-4	Sequence 4, Appl 1
2	4853	100.0	890	US-09-989-687-4	Sequence 4, Appl 1
3	4810	99.1	924	US-10-425-114-39107	Sequence 3, Appl 1
4	3913.5	80.8	905	US-09-918-171A-9	Sequence 9, Appl 1
5	2344	48.3	950	US-09-321-987B-4	Sequence 7, Appl 1
6	2344	48.3	968	US-10-163-316-7	Sequence 7, Appl 1
7	2328.5	48.0	950	US-09-373-658-2	Sequence 2, Appl 1
8	2328.5	48.0	950	US-09-989-687-2	Sequence 2, Appl 1
9	2328.5	48.0	967	US-10-105-929-2	Sequence 2, Appl 1
10	2328.5	48.0	967	US-10-115-286-2	Sequence 2, Appl 1
11	2328.5	48.0	968	US-09-373-658-125	Sequence 125, Appl 1
12	2328	48.0	931	US-09-741-151-4	Sequence 4, Appl 1
13	2320.5	47.8	967	US-09-989-687-126	Sequence 126, Appl 1
14	2274.5	46.9	481	US-09-802-582-8	Sequence 8, Appl 1
15	2274.5	46.9	481	US-10-105-929-8	Sequence 8, Appl 1

16	2274.5	46.9	481	US-10-365-227-8	Sequence 8, Appl 1
17	2146.5	44.2	924	US-10-093-463-28	Sequence 28, Appl 1
18	2138.5	44.1	950	US-09-965-631-4	Sequence 4, Appl 1
19	2138.5	44.1	950	US-09-741-151-2	Sequence 2, Appl 1
20	2128.5	43.9	928	US-10-275-107-59	Sequence 59, Appl 1
21	2115.5	43.6	952	US-10-311-035-11	Sequence 11, Appl 1
22	2103.5	43.3	727	US-09-445-023A-1	Sequence 1, Appl 1
23	2103.5	43.3	727	US-10-097-597-1	Sequence 1, Appl 1
24	2103.5	43.3	727	US-10-097-580-1	Sequence 12, Appl 1
25	2093.5	43.1	727	US-09-445-023A-12	Sequence 12, Appl 1
26	2093.5	43.1	727	US-10-097-597-12	Sequence 12, Appl 1
27	2093.5	43.1	727	US-10-097-580-12	Sequence 12, Appl 1
28	1989.5	41.0	822	US-10-163-316-2	Sequence 57, Appl 1
29	1989	41.0	364	US-09-764-903-57	Sequence 11, Appl 1
30	1842.5	38.0	837	US-10-358-283-11	Sequence 11, Appl 1
31	1842.5	38.0	837	US-10-247-685-2	Sequence 2, Appl 1
32	1841.5	37.9	837	US-09-946-374-317	Sequence 317, Appl 1
33	1841.5	37.9	837	US-10-206-915-352	Sequence 352, Appl 1
34	1841.5	37.9	837	US-10-199-670-352	Sequence 352, Appl 1
35	1841.5	37.9	837	US-10-201-858-352	Sequence 352, Appl 1
36	1841.5	37.9	837	US-10-205-890-352	Sequence 352, Appl 1
37	1841.5	37.9	837	US-10-208-024-352	Sequence 352, Appl 1
38	1841.5	37.9	837	US-10-201-853-352	Sequence 352, Appl 1
39	1841.5	37.9	837	US-10-174-581-352	Sequence 352, Appl 1
40	1841.5	37.9	837	US-10-176-483-352	Sequence 352, Appl 1
41	1841.5	37.9	837	US-10-176-914-352	Sequence 352, Appl 1
42	1841.5	37.9	837	US-10-176-915-352	Sequence 352, Appl 1
43	1841.5	37.9	837	US-10-006-485A-317	Sequence 317, Appl 1
44	1841.5	37.9	837	US-10-013-907A-317	Sequence 317, Appl 1
45	1841.5	37.9	837	US-10-013-907A-317	Sequence 317, Appl 1

# ALIGNMENTS

RESULT 1  
US-09-373-658-4  
Sequence 4, Application US/09373658  
Publicat ion No US20030092900A1  
GENERAL INFORMATION:  
APPLICANT: Iruela-Arispe, Luisa  
APPLICANT: Hastings, Gregg A.  
APPLICANT: Ruben, Steven M.  
APPLICANT: Jonak, Zdenka L.  
APPLICANT: Trulli, Stephen H.  
APPLICANT: Fromwald, James A.  
APPLICANT: Terrett, Jonathan A.  
TITLE OF INVENTION: Mech1 and Mech2 Polynucleotides and Polypeptides  
FILE REFERENCE: 1488.1070006  
CURRENT APPLICATION NUMBER: US/09/373,658  
CURRENT FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 125  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 890  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-373-658-4  
Query Match 100.0%; Score 4853; DB 10; Length 890;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFPAAPAPWLPFLILLILLPLARQAPAPAAAGQASLVVPRTPSSAELIALHUSA 60  
Db 1 MFPAAPAPWLPFLILLILLPLARQAPAPAAAGQASLVVPRTPSSAELIALHUSA 60  
QY 61 FGGGFLRLAPADPSFLAPFEXIERLGGSGRATGCGRGATGCFSGTVNGEPSSLAIVSLC 120  
Db 61 FGGGFLRLAPADPSFLAPFEXIERLGGSGRATGCGRGATGCFSGTVNGEPSSLAIVSLC 120  
QY 121 RGLSGFLLDGEFTTQPGAGSLAOPHRLQWPGAKRPLPRGPEVEVGEQGROR 180  
Db 121 RGLSGFLLDGEFTTQPGAGSLAOPHRLQWPGAKRPLPRGPEVEVGEQGROR 180

```

Db      121  RGLSGSFLDGEFTTQPOGAGSLAOPHRLQRMGPAGARP.LPRPMEVEVTGEQORER 180
Qy      181  GDHODESEESQEEBEGASBPPLGATSTKRFVSEARFVETLLVADASMAFYGADL 240
Db      181  GDHODESEESQEEBEGASBPPLGATSTKRFVSEARFVETLLVADASMAFYGADL 240
Qy      241  QNHILTLMSVARIYKHSIKNSINLMVYKYLIVDEKMGPEVSDNGGLTRNFCNMQR 300
Db      241  QNHILTLMSVARIYKHSIKNSINLMVYKYLIVDEKMGPEVSDNGGLTRNFCNMQR 300
Qy      301  FNQPSDRHPEHYDTAILLTRQNFQGEGLCDTLGVADIGTICDPNKSQSVIEDGLQAAH 360
Db      301  FNQPSDRHPEHYDTAILLTRQNFQGEGLCDTLGVADIGTICDPNKSQSVIEDGLQAAH 360
Qy      361  TLAHELGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLPMSPCSAMYLTLLDG 420
Db      361  TLAHELGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLPMSPCSAMYLTLLDG 420
Qy      421  GHGDCILDPAGALPLPTGLPGRMALYOLDQCQOIFGDFRHCNTSAQVCAQLMCHT 480
Db      421  GHGDCILDPAGALPLPTGLPGRMALYOLDQCQOIFGDFRHCNTSAQVCAQLMCHT 480
Qy      481  DGAEPILCHTKNGSLPMADGTPCGPGLCSBGSCLPEEVEERPKPVVDGMAFWMGES 540
Db      481  DGAEPILCHTKNGSLPMADGTPCGPGLCSBGSCLPEEVEERPKPVVDGMAFWMGES 540
Qy      541  RTGGGVOFSHRECDPEPONGGRYCLGRARAYOSCHTECPDDGKSFREQCEKXNAYN 600
Db      541  RTGGGVOFSHRECDPEPONGGRYCLGRARAYOSCHTECPDDGKSFREQCEKXNAYN 600
Qy      601  YTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFKVFAKVIDGTICGPTLAI CYRG 660
Db      601  YTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFKVFAKVIDGTICGPTLAI CYRG 660
Qy      661  QCVKACDHYVDSPRKLDKGVCGGKNSCRKVSGLPTPTNYGNDIYTI PAGATNIDVK 720
Db      661  QCVKACDHYVDSPRKLDKGVCGGKNSCRKVSGLPTPTNYGNDIYTI PAGATNIDVK 720
Qy      721  QRSHPGVQNDGNLYALKTADGQYLINGNLASAIIEODILVKGITILKXSGSIATLERLOS 780
Db      721  QRSHPGVQNDGNLYALKTADGQYLINGNLASAIIEODILVKGITILKXSGSIATLERLOS 780
Qy      781  RPLPEPLTVOLLTVPEGVPPPKYKTFPVNDVDSMOSSEKERTATNTIIOPLIHAQWVLG 840
Db      781  RPLPEPLTVOLLTVPEGVPPPKYKTFPVNDVDSMOSSEKERTATNTIIOPLIHAQWVLG 840
Qy      841  DMSBCSTCGAGWQRTVECRDPSGQASATCNKALXPEDAKPCESQLCPL 890
Db      841  DMSBCSTCGAGWQRTVECRDPSGQASATCNKALXPEDAKPCESQLCPL 890

```

## RESULT 2

```

US-09-989-687-4
; Sequence 4, Application US/09989687
; Publication No. US20040002449A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 890
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-989-687-4

```

```

Query Match      100.0%; Score 4853; DB 11; Length 890;
Best Local Similarity 100.0%; Pred No. 0;
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MFPAPAPRWLPFLILLLLPLARGAPAPPAAGQASELVPTRLPGSAGELALHUSA 60
Db      1  MFPAPAPRWLPFLILLLLPLARGAPAPPAAGQASELVPTRLPGSAGELALHUSA 60
Qy      61  FKGGLVRLAPDSDFLAEPFKIERIGSGGATGGERGLRGCFPSGTVNGEPESLAAVSLC 120
Db      61  FKGGLVRLAPDSDFLAEPFKIERIGSGGATGGERGLRGCFPSGTVNGEPESLAAVSLC 120
Qy      121  RGLSGSFLDGEFTTQPOGAGSLAOPHRLQRMGPAGARP.LPRPMEVEVTGEQORER 180
Db      121  RGLSGSFLDGEFTTQPOGAGSLAOPHRLQRMGPAGARP.LPRPMEVEVTGEQORER 180
Qy      181  GDHODESEESQEEBEGASBPPLGATSTKRFVSEARFVETLLVADASMAFYGADL 240
Db      181  GDHODESEESQEEBEGASBPPLGATSTKRFVSEARFVETLLVADASMAFYGADL 240
Qy      241  QNHILTLMSVARIYKHSIKNSINLMVYKYLIVDEKMGPEVSDNGGLTRNFCNMQR 300
Db      241  QNHILTLMSVARIYKHSIKNSINLMVYKYLIVDEKMGPEVSDNGGLTRNFCNMQR 300
Qy      301  FNQPSDRHPEHYDTAILLTRQNFQGEGLCDTLGVADIGTICDPNKSQSVIEDGLQAAH 360
Db      301  FNQPSDRHPEHYDTAILLTRQNFQGEGLCDTLGVADIGTICDPNKSQSVIEDGLQAAH 360
Qy      361  TLAHELGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLPMSPCSAMYLTLLDG 420
Db      361  TLAHELGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLPMSPCSAMYLTLLDG 420
Qy      421  GHGDCILDPAGALPLPTGLPGRMALYOLDQCQOIFGDFRHCNTSAQVCAQLMCHT 480
Db      421  GHGDCILDPAGALPLPTGLPGRMALYOLDQCQOIFGDFRHCNTSAQVCAQLMCHT 480
Qy      481  DGAEPILCHTKNGSLPMADGTPCGPGLCSBGSCLPEEVEERPKPVVDGMAFWMGES 540
Db      481  DGAEPILCHTKNGSLPMADGTPCGPGLCSBGSCLPEEVEERPKPVVDGMAFWMGES 540
Qy      541  RTGGGVOFSHRECDPEPONGGRYCLGRARAYOSCHTECPDDGKSFREQCEKXNAYN 600
Db      541  RTGGGVOFSHRECDPEPONGGRYCLGRARAYOSCHTECPDDGKSFREQCEKXNAYN 600
Qy      601  YTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFKVFAKVIDGTICGPTLAI CYRG 660
Db      601  YTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFKVFAKVIDGTICGPTLAI CYRG 660
Qy      661  QCVKACDHYVDSPRKLDKGVCGGKNSCRKVSGLPTPTNYGNDIYTI PAGATNIDVK 720
Db      661  QCVKACDHYVDSPRKLDKGVCGGKNSCRKVSGLPTPTNYGNDIYTI PAGATNIDVK 720
Qy      721  QRSHPGVQNDGNLYALKTADGQYLINGNLASAIIEODILVKGITILKXSGSIATLERLOS 780
Db      721  QRSHPGVQNDGNLYALKTADGQYLINGNLASAIIEODILVKGITILKXSGSIATLERLOS 780
Qy      781  RPLPEPLTVOLLTVPEGVPPPKYKTFPVNDVDSMOSSEKERTATNTIIOPLIHAQWVLG 840
Db      781  RPLPEPLTVOLLTVPEGVPPPKYKTFPVNDVDSMOSSEKERTATNTIIOPLIHAQWVLG 840
Qy      841  DMSBCSTCGAGWQRTVECRDPSGQASATCNKALXPEDAKPCESQLCPL 890
Db      841  DMSBCSTCGAGWQRTVECRDPSGQASATCNKALXPEDAKPCESQLCPL 890

```

## RESULT 3

```

US-10-425-114-39107
; Sequence 39107, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

```

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313) B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ. ID NOS: 73128  
SEQ. ID NO: 39107  
LENGTH: 924  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB4654-025-F6\_P1.pcp  
US-10-425-114-39107

Query Match 99.1%; Score 4810; DB 12; Length 924;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 884; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

1 MFPAAPRMLPFLLLLLLLPLAAGAPAPPAAGGASELVPTRLPGSAGELALHLSA 60  
37 MLPAAPRMP-LLLLLLPLAAGAPAPPAAGGASELVPTRLPGSAGELALHLSA 94  
61 FGKGFVRLAPDSDFLAPFEKIERLGSGGATGGERLRCGCFSGTVNGEPESLAIVSLC 120  
95 FGKGFVRLAPDSDFLAPFEKIERLGSGGATGGERLRCGCFSGTVNGEPESLAIVSLC 154  
121 RGLSGSFLDGEETTPQAGAGSLAOPHRLORNGAPAPLPRGEMETEGEGORER 180  
155 RGLSGSFLDGEETTPQAGAGSLAOPHRLORNGAPAPLPRGEMETEGEGORER 214  
181 GDHODESEESQEEBAGASPPPLGATSTRKFSVSEAPFVETLLVADASMAAFYADL 240  
215 GDHODESEESQEEBAGASPPPLGATSTRKFSVSEAPFVETLLVADASMAAFYADL 274  
241 QNHLLTMSVAARIYKPSISINLMVVKYLVEDKMGPEVSDNGQLTLRFQWQR 300  
275 QNHLLTMSVAARIYKPSISINLMVVKYLVEDKMGPEVSDNGQLTLRFQWQR 334  
301 FNGSDRPEHYDAIILTRONFCQEGELCDTLGVADIGTICDPMKCSVLEDEGLQAAH 360  
335 FNGSDRPEHYDAIILTRONFCQEGELCDTLGVADIGTICDPMKCSVLEDEGLQAAH 394  
361 TLAEHLGVLNMPHDSKPCRTLFGPMKHHVMAPLFHLNQLTPMSPCSAMYTELLDG 420  
395 TLAEHLGVLNMPHDSKPCRTLFGPMKHHVMAPLFHLNQLTPMSPCSAMYTELLDG 454  
421 GHGDCLLAPGAALPLPTGLPGRMALYQLDQCCROIFGPDFRHCENISAQVCAQLWCHT 480  
455 GHGDCLLAPGAALPLPTGLPGRMALYQLDQCCROIFGPDFRHCENISAQVCAQLWCHT 514  
481 DGAEBLCHTKNGSLPWADGTPCGPHLCSESSCLPEEVEVERPKPVVDGMAFWGPGSCS 540  
515 DGAEBLCHTKNGSLPWADGTPCGPHLCSESSCLPEEVEVERPKPVVDGMAFWGPGSCS 574  
541 RCGGGVQFSEHCECDPEPONGGRYCLGRKAYGSCHEECPPDCKSPRECCCKYANN 600  
575 RCGGGVQFSEHCECDPEPONGGRYCLGRKAYGSCHEECPPDCKSPRECCCKYANN 634  
601 YTDMDGNLLQWPKTAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLGSEPTLAIYVRG 660  
635 YTDMDGNLLQWPKTAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLGSEPTLAIYVRG 694  
661 QCVKAGCHVAVDSKRLDKCVCGGKNGSCRVKSLPTNYGNDYITIRAGATNIDYK 720  
695 QCVKAGCHVAVDSKRLDKCVCGGKNGSCRVKSLPTNYGNDYITIRAGATNIDYK 754  
721 QSHSPVQNDGNVLAALKTADQYTLNGNLTAISAIBODILVKGTILKYSGSITATLERSF 780  
755 QSHSPVQNDGNVLAALKTADQYTLNGNLTAISAIBODILVKGTILKYSGSITATLERSF 814  
781 RPLPELITVQLTVPESEVPPKRYTFVPPNDVDSMOSKERRATNTIIOPLIHAQWVLG 840  
815 RPLPELITVQLTVPESEVPPKRYTFVPPNDVDSMOSKERRATNTIIOPLIHAQWVLG 874

QY 841 DMSRCSSTCGAMORITVECRDPGASATCNKALKEPDAPCEBQUCPL 890  
DB 875 DMSRCSSTCGAMORITVECRDPGASATCNKALKEPDAPCEBQUCPL 924

RESULT 4  
US-09-918-171A-9  
Sequence 9, Application US/09916171A  
Patent No. US20020110894A1  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Titta L.  
APPLICANT: Hirohata, Satoru  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/04193  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 09/365,364  
PRIOR FILING DATE: 1999-08-06  
NUMBER OF SEQ. ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ. ID NO: 9  
LENGTH: 905  
TYPE: PRT  
ORGANISM: Mus musculus ADAMTS-8  
US-09-918-171A-9

Query Match 80.8%; Score 3919.5; DB 9; Length 905;  
Best Local Similarity 79.2%; Pred. No. 5,4e-303;  
Matches 717; Conservative 74; Mismatches 99; Indels 15; Gaps 4;

1 MFPAAPRMLPFLLLLLLLPLAAGAPAPPAAGGASELVPTRLPGSAGELALHLSA 59  
1 MRDPTTGWPLLLLLLLQLEPPPLVCAAPGPGTGAQASELVPTRLPGSAGELALHLSA 60  
60 AFGKGFVRLAPDSDFLAPFEKIERLGSGGATGGERLRCGCFSGTVNGEPESLAIVSL 119  
61 AFGKGFVRLAPDSDFLAPFEKIERLGSGGATGGERLRCGCFSGTVNGEPESLAIVSL 120  
120 CRGLSGSFLDGEETTPQAGAGSLAOPHRLORNGAPAPLPRGEMETEGEGORER 168  
121 VAGMSGSFLDGEETTPQAGAGSLAOPHRLORNGAPAPLPRGEMETEGEGORER 180  
169 EVATGEGQOEGRDHOODESEESQEEBAGASPPPLGATSTRKFSVSEAPFVETLL 226  
181 EVATGEGQOEGRDHOODESEESQEEBAGASPPPLGATSTRKFSVSEAPFVETLL 240  
227 VADASMAAFYADLQNHLLTMSVAARIYKPSISINLMVVKYLVEDKMGPEVSDN 286  
241 VADASMAAFYADLQNHLLTMSVAARIYKPSISINLMVVKYLVEDKMGPEVSDN 300  
287 GGLTLNFCWQRRFNPQSDRPEHYDAIILTRONFCQEGELCDTLGVADIGTICDPMK 346  
301 GGLTLNFCWQRRFNPQSDRPEHYDAIILTRONFCQEGELCDTLGVADIGTICDPMK 360  
347 SCSEVIEDGLOAAHTLAHELGHVLSMPHDSKPCRTLFGPMKHHVMAPLFHLNQLTPM 406  
361 SCSEVIEDGLOAAHTLAHELGHVLSMPHDSKPCRTLFGPMKHHVMAPLFHLNQLTPM 420  
407 SPSCAMVLTBLDGGHDCCLDAPGAALPLPTGLPGRMALYQLDQCCROIFGPDFRHC 466  
421 SPSCAMVLTBLDGGHDCCLDAPGAALPLPTGLPGRMALYQLDQCCROIFGPDFRHC 480  
467 TSAQDVCAQLMC-HTDGAEBLCHTKNGSLPWADGTPCGPHLCSESSCLPEEVEVERPK 525  
481 TSAQDVCAQLMC-HTDGAEBLCHTKNGSLPWADGTPCGPHLCSESSCLPEEVEVERPK 540  
526 VDGMAFWGPGSCSRKCPGCGGVQFSEHCECDPEPONGGRYCLGRKAYGSCHEECPPD 585  
541 VDGMAFWGPGSCSRKCPGCGGVQFSEHCECDPEPONGGRYCLGRKAYGSCHEECPPD 600  
586 KSPRECCCKYANNYTDMDGNLLQWPKTAGVSPDRCKLFCRARGSEFKVFEAKVID 645

Fri May 7 12:18:02 2004

us-09-989-687-4\_1.rapb

Page 4

Db	601	KSFEQCCCEXVNYMHDLGNFLQVWVPKSGVSPDRCKLFCRARGHSEKVFPEAVIID	666
Qy	646	GTLCGPEPTLIAICVGGCCVYKAGCDHYVDSPEKLDKCGCGGKENSRCRXYSGSLTPNKGYN	705
Db	661	GTLCGPPTLSTCYRGCCVYKAGCDHYVNSPKDKDGCVCGKETAACKLSGSEFTPTSTGYN	720
Qy	706	DIYTIIPAGANINIDVKQSGSHFGVNDGNYALKTADQGYILNGNLIAISAIIBODIIVKGTIL	765
Db	721	DIYTIIPAGANINIDVKQSHFGVNDGNSYALKTANGQYILNGNLIAISAIIBODIIVKGTIL	780
Qy	766	KYSGSIAITLERLOSFPLEPEPTLVOLLTPGGEVFPKKYVTFEVENDVDVFGMSGKERAT	825
Db	781	KYSGSMATLERLOSFPALPEPTLVOLLTPGGEVFPKKYVTFEVENDMDVFGVMSKERAT	840
Qy	826	TNIIQPLHLOWTLGDMSECCSSTCGAGMORFVTECRDSCGASATCNAALPEPAKCCES	885
Db	841	TNIIQPLPSALEWTLGDMSECCSPTCRDSKMRVTECRDSCGASDPCDPAALPEPAKCCGS	900
Qy	886	QLGCL 890	
Db	901	QpCZL 905	

```

RESULT 5
US-09-331-987B-4
Sequence 4, Application US/09321987B
Patent No. US20020102210A1
GENERAL INFORMATION:
APPLICANT: Kimble, Judith E
APPLICANT: Bielecki, Robert H
TITLE OF INVENTION: Agent and Method for Modulating Cell Migration
FILE REFERENCE: 960296, 95386
CURRENT APPLICATION NUMBER: US/09/321,987B
PRIORITY FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 60/087,170
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 60/129,023
PRIOR FILING DATE: 1999-04-13
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 950
TYPE: PRT
ORGANISM: Murine
US-09-331-987B-4

```

```

Query Matched          48.3%; Score 2344; DB 9; Length 950;
Best Local Similarity 49.5%; Pred. No. 29-177;
Matches 457; Conservative 143; Mismatches 241; Indels 78; Gaps 20

QY      14 LLLLLLLLP LARGAPARPAGQASSELVFT--RLPG-SAGELALHLSFGKGYRLA 70
Db      21 LLLASTITMLICRGAGHPTE--EDELIVLPSLRRAGHDSITRYRLIDFGQQLTKQ 78

QY      71 PDGSFLAPFKETLERGGSGRATGGR-----GLRGCFSGTMYNBPBSLAVISCRL 123
Db      79 PDGSGFLAPGFTLTQTV---GRSPGSAQGLDPTGCLACFTSGYNADPGSAALSLCEGV 135

QY      124 SGSFLLDGEEFTTPOGAGSLAQRLQRMGPAGAPRLPRGEWEYETEG---ORQE 179
Db      136 RGAFLYLGEBEFTIOPAP-----GVATERLAPVPEESANPQHILRRR 181

QY      180 RGD-----HODEEESQEEAEAGAEPPRPGLA-----TSRKRPVLSA 219
Db      182 RSGGKAGCGWMDDETLPTSDSRPESQNTRNQWRYVRLDPPDAGKPSGPGSIRKRRVSSP 241

QY      220 RPYETLLVADASAAAFYGADLQNTLLITMVSAAIYKHPISIKSNINIMVKKVLIVEDKK 279
Db      242 RYVETMLVADQSMADPHGSGGLKHYLLTFVSAAFYKHPISIRNISTLVVKKILVIEEK 301

QY      280 GPVPSDNGGLLRNCCNNQRRFNQSSDHPHHYDTALLLRQVCCGEGLCPTLGVADIG 339
Db      ||||:|||||||:::||||:||||:||||:||||:||||:||||:||||:||||:

```

```

Db      302  GREVTNNALTLTNFECNMWQKHNSSBDDPEHYDAILFTQRQDCSGH-TCDDTGMADV 360
Qy      340  TIDCPKSSGVSVIEDEGLQAAHTLAEHLVSMPEHDSKCPOTRLFGMGKNHWAPLFLVH 399
Db      361  TVCDPERSGSVIEDDLQAFHTTAEHLGVFEMPHDDKQKASLNGVTGDSHMASTLSS 420
Qy      400  INQTLWSPSCSAWYITELLDDGHBDCLLDARPAALPLETGLPGMALYOLDQOCROITGP 459
Db      421  LDHSQWSPSCSAWYATSLFDLNGHSGCLMDKQNPRIKLPSPDJG--TLYDANRQOCFTFGE 478
Qy      460  DFHCNENTSADQVCAQLWC-HTDGAELPLCHTKNSLPLPADGTPCGPGHLCESGCLPEEE 518
Db      479  ESKHCD--DASTCTTLMCTGSHGGLVQTRK--FPMADSTSGEGKWCYSGKCVAKTD 534
Qy      519  VERPKVVDGMAWPMWPMWBCSRTGSGVQFSHEKCPDEBQNGRAYCLGRPAKYOSSHT 578
Db      535  MKGFAPFVSGSGPMWGMWBCSRTGSGGVQYTWMECONPVPKNGKCEBGRVAYSCNI 594
Qy      579  EEGCP--DGSFPEOQCEKNNAYNYTDMQGN--LLQWPKYAGVSPRRCXLYFGARSRSE 635
Db      595  EDGPDNNGKTFEEBQCEAHNERSKSF--GNELPYEMTPKAYGVSXPORCKLTCAKRTGY 653
Qy      636  FKYEFAKVIDGTLCSBETLALCTRGQCCYKAGCNDHVDSPKLDCXCYCGSGKNSCKRVSG 695
Db      654  FFLQKRVYVDGPRCCBPDSTSVCCQCYKACDRLIDSKKFFDCCGVCGGNSCKCKMSG 713
Qy      696  SLTPYVGYNDIVTIPAGATNIDVQKSHSPVQNDGNVLTAKTAGOYLNLNGNATAISAE 755
Db      714  IYVSTRPGVHDIVTIPAGATNIEVGHNRQKSRNNGSLAIRADGTYILNGNPTSTLE 773
Qy      756  QDLVAVGTLTKKSGSLATLERLQSGRPLREPLTYQLTYBEEVFRPKYKTFPRNDVDF 815
Db      774  QDLTYVGTVALRVSGSSAALERTSRSPLEKEPLTYQLVNV--GHALRPKIKTYFN----- 826
Qy      816  SMQSKERATYIIIPLLHAWQVLGDMSECSSTGAGNMORTVECRDPSCGASATCNKAL 875
Db      827  ---KKKTESFNALPTF--SEWIVBEMGECKTSGSGMQRVVQCRDINGHAPASECAKEY 880
Qy      876  KPEDAKPCESQCLP 889
Db      881  KPASTRPCADLP 894

```

```

RESULT 6
US-10-163-316-7
Sequence 7: Application US/10163316
Publication No. US20020197703A1
GENERAL INFORMATION:
APPLICANT: Kapeller-Libermann, Rosana
TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses
TITLE OF INVENTION: Therefore
FILE REFERENCE: MP101-025P1PM
CURRENT APPLICATION NUMBER: US/10/163,316
CURRENT FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/297,863
PRIOR FILING DATE: 2001-06-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO. 7
LENGTH: 968
TYPE: PRT
ORGANISM: Mus musculus
US-10-163-316-7

```

[illegible]

```

Db      96 PDGSLFAPGFTLTQTV---GRSPGSEAOHLDPDTDLACHFSGVNDPQSAALSLCEVY 152
Qy      124 SSGSLDGEFEFTTPOGAGGSLAOPHRLQWGPAGARPLPRGEMEVETGEG---QROE 179
      153 RGAFTYLOGEEFTTPOAP-----GVATERLAIVAEESASRPOPHILRRR 198
Qy      180 RGD-----HOEDSEESOEAEAGSEPPPLGA-----TSRTKRVSA 219
      199 RGSAGACGVWDEDTLPTSDSRPESQNTNQMVPDPTPODAKPSGPGIRKRVSSP 258
Qy      220 RFVETTLVADASMAAFGADLONHILTMGVAARIYGHPIKASINIMVYKLVIEDKX 279
      259 RYETMLVADOSNADFFSGGKHYLLTFVAARFYHPIRISISIVVKKIIVIEBK 318
Qy      280 GREVSNDGJLTNRNFCNWRFRNOPSDBRPHYDTAILTRONFCQOEGLCDTLGVADIG 339
      319 GREVTSNAALTTRNFCNWKONPSDRDEHYDTAILFTRODLGSH--TCDTLGMADVG 377
Qy      340 TICDPNKSCTVIEDEGLQAAHTLAHELGHVLSMHPDQSKPCTRLFGMGKHVMAPLFVH 399
      378 TVCDPSSCSYIEEDGLQAAFTTAHELGHVFNPHDQKQASLNGVTDGSHLMAASLS 437
      400 LNOTLPMSPCSAMYLTEILDGSHDCLLDAPGALPLPTGLPGRMALYOLDQOCROIFGP 459
      438 LDHSGPMSPCSAVWVTSFLDNGHGECLMDKQNPILKPSDLLP--TLYDANRQCCFTFGE 495
Qy      460 DFRHCNTSAQDVCAQIMC-HTDGAEPULCHTKNGSLPMADGTFCGPHLCSBSCLEPBE 518
      496 ESKHCPD--AASTCTTLMCTGTSGLLYCQTKH--FPMADGTSCGEGKWCVSCKVKNKD 551
Qy      519 VERKPYVDGGMAPMGWEGCSRTCGGVOFSHRECKDPBPONGRXYCLGRARAKYOSCHT 578
      552 MKHATVPHSGWGMWGMGDCSRICGSGVOYTMRECONPYPKQKXCEBKRYRYSN 611
Qy      579 EECBP-DGKSFREOCCERKNANVTMDGN--LIQVWPXYAGVSPDRDCLFCRARGSE 635
      612 EDCDNNKCTFREBOCEAHNEFSKASF-GNEPVEWMPKTAGVSPKDRCLTLEAKIGY 670
Qy      636 FKVEEAKYIDTLCGPETLAIQVAGCCVYKAGCDHVVSPPRLDKCGVCGKNGSCRYVG 695
      671 FFIQOPKVVDTGSPSPSTSVCGQGVKACGCDRIIDSKKFKPCGVCGNGSGCTCKMSG 730
Qy      696 SLFTPNVGYNDIVITIPAGATNIDVYKORSHPGVQNDGNVYALKTADGOYLINGMLAISALE 755
      731 IVTSTRPYHDIVITIPAGATNIEVYKHNRGSRNNGSFLAIRADGYIILNGVFTSTLE 790
Qy      756 QDILVKGITLKYSGIATLERLOSFRPLPEPLTVQULTVPGSEVFPKVKYKTFVPRNDVF 815
      791 QDILVKGITLKYSGIATLERLOSFRPLPEPLTVQULTVPGSEVFPKVKYKTFVPRNDVF 843
Qy      816 SMQSKKATNTNIQPLHQAQVLDGMECSSTCGAGMORVTECRDPDSGOASATCNKAL 875
      844 ----KKKTESNAPLPTF--SEWVIEHMEGCSKTCGSGMQRVYVCGRIINGHPSACEKVEY 897
Qy      876 KPEDAKPCESQLCP 889
      898 KPASTRPCADLPCP 911

```

```

RESULT 7
US-09-373-658-2
; Sequence 2, Application US/09373658
; Publication No. US20030092900A1
; GENERAL INFORMATION:
; APPLICANT: Inuela-Aripe, Luisa
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Zdenka, Zdenka L.
; APPLICANT: Trull, Stephen H.
; APPLICANT: Fromwald, James A.
; APPLICANT: Terrett, Jonathan A.
; TITLE OF INVENTION: Methyl and Methyl Polynucleotides and Polypeptides

```

```

; FILE REFERENCE: 1488,1070006
; CURRENT APPLICATION NUMBER: US/09/373,658
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 950
; TYPES: PRT
; ORGANISM: Homo sapiens
US-09-373-658-2

Query Match      48.0%; Score 2328.5; DB 10; Length 950;
Best Local Similarity 49.7%; Pred. No. 3,4e-16;
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

Qy      11 LPFLILLILLILLPLARGPAPAPAGGASSELVVP--TELPSAGEIALHLSAFGKGYLR 68
      16 VFTLLLLAALLAAS-DLGRPSE--EDBELVPELEKAPGH-GTTRLRLLAPQQLDLE 71
Qy      69 LAPDSFLAPEKIERLQSGSRATG-----ERGLRGCFSGTYNGEPESLAASLCRCL 123
      72 LRPDSSFLAPFTLQNV---GRKSGSERPLPETDLACHFYAGTYGDSAAALSLCEGV 128
Qy      124 SSGSLDGEFEFTTPOGAGGSLAOPHRLQWGPAGARPL-----RGPEMEV---ET 172
      129 RGAFTYLOGEEFTTPOAPAS-----ERLATAPEKRPAPLQFILRRNRQDVGCTGV 183
Qy      173 GEGORGEHDHDEDESESOEAEAGASEPP-----PPLGATS-RTKRFVSEARFV 222
      184 VDERPRTGKATEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 243
Qy      223 ETLVADASMAAFYADLONHILTMGVAARIYKHPIKASINIMVYKLVIEDKMBE 282
      244 ETLVADOSNADFFSGGKHYLLTFVAARLYKHPIRISISIVVKKIIVIEBKQPE 303
Qy      283 VSDNGGLTLNRNFCNWRFRNOPSDBRPHYDTAILTRONFCQOEGLCDTLGVADIGTIC 342
      304 VTSNALTNRNFCNWKONPSDRDEHYDTAILFTRODLGSH--TCDTLGMADVGTVIC 362
Qy      343 DNKSCSVIIEDEGLQAAHTLAHELGHVLSMHPDQSKPCTRLFGMGKHVMAPLFVHLNQ 402
      363 DSRSCSVIIEDEGLQAAFTTAHELGHVFNPHDQKQASLNGVTDGSHLMAASLSNDH 422
Qy      403 TLPMSPCSAMYLTEILDGSHDCLLDAPGALPLPTGLPGRMALYOLDQOCROIFGPDR 462
      423 SQPMSPCSAVWVTSFLDNGHGECLMDKQNPILKPSDLLP--TLYDANRQCCFTFGEBSK 480
Qy      463 HCENTSAQDVCAQIMC-HTDGAEPULCHTKNGSLPMADGTFCGPHLCSBSCLEPBEVER 521
      481 HCPD--AASTCTTLMCTGTSGLLYCQTKH--FPMADGTSCGEGKWCVSCKVKNKD 536
Qy      522 PKPVVDGGMAPMGWEGCSRTCGGVOFSHRECKDPBPONGRXYCLGRARAKYOSCHTBC 581
      537 PTPFHSGWGMWGMGDCSRICGSGVOYTMRECONPYPKQKXCEBKRYRYSNCLBEC 596
Qy      582 PP-DGKSFREOCCERKNANVTMD--DGNLQVWPXYAGVSPDRDCLFCRARGSEFKVF 639
      597 PDNNGKTFREBOCEAHNEFSKASF-GNEPVEWMPKTAGVSPKDRCLTLEAKIGYFV 656
Qy      640 EAKYIDTLCGPETLAIQVAGCCVYKAGCDHVVSPPRLDKCGVCGKNGSCRYVGSULTP 699
      657 QPKVVDGTPCSPSTSVCGQGVKACGCDRIIDSKKFKPCGVCGNGSGCTCKKISGSVTS 716
Qy      700 TNYGVNDIVITIPAGATNIDVYKORSHPGVQNDGNVYALKTADGOYLINGMLAISALEQDIL 759
      717 AKGCHDITITIPAGATNIEVYKHNRGSRNNGSFLAIRADGYIILNGVFTSTLEQDIL 776
Qy      760 VKGTILKYSGIATLERLOSFRPLPEPLTVQULTVPGSEVFPKVKYKTFVPRNDVPSMS 819
      777 YKGVVLYRSGSSAALRIRISFSLKEPLITQVLTV--GNALRPIKIXTYFV----- 825
Qy      820 SKERATNTNIQPLHQAQVLDGMECSSTCGAGMORVTECRDPDSGOASATCNKALKPED 879

```

Fri May 7 12:18:02 2004

us-09-989-687-4\_1.rapb

Page 6

Db 826 KKKKESFNAL-PTFSA-WIIEWGECSCSKELGMORRLVECRDINGOPASECAKEVFPAS 883  
QY 880 AKPCESQLCP 889  
Db 884 TRPCADHPCP 893

RESULT 8

US-09-989-687-2  
; Sequence 2, Application US/09989667  
; Publication No. US20040002449A1  
; GENERAL INFORMATION:  
; APPLICANT: Hastings, Gregg A.  
; APPLICANT: Ruben, Steven M.  
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides  
; FILE REFERENCE: 1488.107000D  
; CURRENT APPLICATION NUMBER: US/09/989,687  
; CURRENT FILING DATE: 2001-11-21  
; NUMBER OF SEQ ID NOS: 126  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 950  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-989-687-2

Query Match 48.0%; Score 2328.5; DB 11; Length 950;  
Best Local Similarity 49.7%; Pred. No. 3,4e-176;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

QY 11 LPTLLILLILLPLLAGAPAPPAAGQASEIWP--TRLPSAGELALHLSAFKGFVLR 68  
Db 16 VPTLLLLAALLAVS-DALGRPSE--EDEELVPELEBPQH-GTTRLRLHAFDQDLLE 71  
QY 69 LAPDSFLAPEFKIRLGGSGRATG-----ERGLRGCFPGTGVNGEPESLAASLORGL 123  
Db 72 LRPDSSFLAPGFTLQNV--GRKSGSETPLPETDLAHCFYSGTGVNDSSAALSLCEGV 128  
QY 124 GGSFLDGEFTIOPQAGAGSLAQPHRLQWGPAGARLP-----RGPEWV---ET 172  
Db 129 RGAFTILGAYFIQPLPAPS-----ERLATAAPGKPPAPLQPHLLRNQGVGTGV 183  
QY 173 GEGORGERGDHODESEESQEEAEAGASEPP-----PPLGATS-RTRKRVSEARFV 222  
Db 184 VDDEPRPTGKATEDEDEGEDEGPQWSPQDPALQGVQPTGTGSTRKRVSSHRVY 243  
QY 223 ETLIVADAMAFYAGDLQNHILTLMSVARIYKHSIKNSINLMVWVILVDEKMGPE 282  
Db 244 ETLIVADQSMAEFHSGLKHVLLTLFVAARLYKHSIRNSVSLVAVKILVIHDEQKPE 303  
QY 283 VSDNGGLTRNFCNWRFPQSDRHPENHYDTAILLTRQNFQCGQGLCDTLGVADIGTIC 342  
Db 304 VTSNALLTRNFCNWKQNPSPDRDAEHYDTAILFTRQDLGSGQ--TCDTLGMADVGTVC 362  
QY 343 DPKKCSVTEDEGLQAHTLAHLEHLYLSMHPDSDKPCRRLLFGPMGKHVNAVPLFVHNO 402  
Db 363 DPKRSQSVTEDEGLQAHTLAHLEHLYLSMHPDSDKPCRRLLFGPMGKHVNAVPLFVHNO 422  
QY 403 TLPWSCSAMVLTLLDGGHGDCLDAPGALPLPTGLFGPMALYQDLQCGRQIFGPDFR 462  
Db 423 SQWMSCSAMVLTSLFDNCHGEGDKMPQNPQLPDPDLP--TSYDARQCCFTGEDSK 480  
QY 463 HCPNTSAQDVCAQJLC-ATDGAEPFLCHTKNGSLPNAADTGPCCGPHLCSGSGCLPEEER 521  
Db 481 HCPD--AATCSTLWCTGTSGVLYQPTKH--FPWADGTSCEBGMKINGAKVNTKDKH 536  
QY 522 PKVYVNGWAPMGPMGECRTCGGVQFHSRECKDPEPONGRKYLCGRATYQSGHTEBC 581  
Db 537 FDFPFHSGMGMKPMDCRTCGGVQYIMRECDNVFPNGGKVGEGKRVYRSQNLBDC 596  
QY 582 PP-DGKSFREQCEKTNAYITDM-DGNILQWPKYAGVSPDRCLLFRAGRSEBFKYF 639  
Db 597 PDNNGKTRFEEQCEAHNEFSKASFGSGPAVEWIPKYAGVSPDRCLLQCAKGIYGFVL 656

QY 640 EAKVIDGTLGCPETALICVRGQCYAKGCDHVVDSPRKLDKGCVGKSGKSKYSGSLP 699  
Db 657 QPKVVDGTPCSPDSTVCGQCVAVAGCDRIIDSKKFPDKGCVGKSGKSKYSGSLP 716  
QY 700 TNYGYNDIYTIIPAGATNIDVKORSHPGVQNDGNVLAITKTAQOYLNLGNLAISAEODIL 759  
Db 717 AKPGYHIDITITPGATNIEVQRNQRGSRNNSFLAIAADGTIILNDYTLSTLEODIM 776  
QY 760 VGTILKTSGSTATTIERLOSFPPLPEPLTVQLLITPGEVFPKXKTYTFVNDVDFSMQS 819  
Db 777 YKGVVLRYSGSAALEIRSFPLEPILTVLTV-GNALRKLKTYTFV----- 825  
QY 820 SKERATNTIOPLLHAOWVLGDMSECSSTCGAGMORRTVECRDPGQASATCNKALKEPD 879  
Db 826 KKKKESFNAL-PTFSA-WIIEWGECSCSKELGMORRLVECRDINGOPASECAKEVFPAS 883  
QY 880 AKPCESQLCP 889  
Db 884 TRPCADHPCP 893

RESULT 9  
US-10-105-929-2  
; Sequence 2, Application US/10105929  
; Publication No. US20020137142A1  
; GENERAL INFORMATION:  
; APPLICANT: Goodheart, Andrew D.J.  
; APPLICANT: Holtzman, Douglas A.  
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
; FILE REFERENCE: 09404/041001  
; CURRENT APPLICATION NUMBER: US/10/105,929  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 967  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-105-929-2

Query Match 48.0%; Score 2328.5; DB 13; Length 967;  
Best Local Similarity 49.7%; Pred. No. 3,4e-176;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

QY 11 LPTLLILLILLPLLAGAPAPPAAGQASEIWP--TRLPSAGELALHLSAFKGFVLR 68  
Db 33 VPTLLLLAALLAVS-DALGRPSE--EDEELVPELEBPQH-GTTRLRLHAFDQDLLE 88  
QY 69 LAPDSFLAPEFKIRLGGSGRATG-----ERGLRGCFPGTGVNGEPESLAASLORGL 123  
Db 89 LRPDSSFLAPGFTLQNV--GRKSGSETPLPETDLAHCFYSGTGVNDSSAALSLCEGV 145  
QY 124 GGSFLDGEFTIOPQAGAGSLAQPHRLQWGPAGARLP-----RGPEWV---ET 172  
Db 146 RGAFTILGAYFIQPLPAPS-----ERLATAAPGKPPAPLQPHLLRNQGVGTGV 200  
QY 173 GEGORGERGDHODESEESQEEAEAGASEPP-----PPLGATS-RTRKRVSEARFV 222  
Db 201 VDDEPRPTGKATEDEDEGEDEGPQWSPQDPALQGVQPTGTGSTRKRVSSHRVY 260  
QY 223 ETLIVADAMAFYAGDLQNHILTLMSVARIYKHSIKNSINLMVWVILVDEKMGPE 282  
Db 261 ETLIVADQSMAEFHSGLKHVLLTLFVAARLYKHSIRNSVSLVAVKILVIHDEQKPE 320  
QY 283 VSDNGGLTRNFCNWRFPQSDRHPENHYDTAILLTRQNFQCGQGLCDTLGVADIGTIC 342



Fri May 7 12:18:02 2004

us-09-989-687-4\_1.rapb

321 VTSNAALTNRFCWQKXHPSPDRDAEHYDIALFTRODLGSGQ--TCDTLGMADVGTVC 379  
343 DPNKSCSVIEDBGLQAAHTLAHEIGHVLSMPHDSKPCTRLFGPMKHHVMAPLFVHLNQ 402  
380 DPNKSCSVIEDDGLQAAFTTAHEIGHVFNMPHDDAKOCASLNGNODSHMMAKMLNLDH 439  
403 TLPMSPCSAMVLTLLDGGHDCILDAFGALPLPTGLPGRMALYQDDQCRQIFGDFR 462  
440 SQPWSPCSAMVLTSLDNGHGBCLMDKPNPQLPGLPG--TSYANRQCCGTFGBDCK 497  
463 HCPNTSADVCAQIMC-HTDGAEPILCHTKNGSLPMWADGTPCGPGLHCSESSCLPEEVER 521  
498 HCPD--AASCTSLMCTGTSGGVAVLCQTKH--FPWADGTSCEGKMCINKCVAKTDKRA 553  
522 PKPVVDGGMAMPWGMGESSRTCCGGVQFNSHRECKDEPONGGRYCLGRARAKYOSCHTEEC 581  
554 FDTFFHSGMWGMWGMGDCSRTCCGGVQYTMRECDNVPKNGGKICBGRVRYRSCNLEDC 613  
582 PP-DGKSFREOQCEKXNAVNTDM-DGNLLQWPKYAGVSPRDRCKLFCRARGSEPKYF 639  
614 PDNNGKTFREOCEAHNBFSSKASFGSGPAVEMIPKAVGSPKRCKLICQAKIGYFVL 673  
640 EAKYIDGTLGPEPTLAICVRGQCVKAGCDHVVDSPKLDKCGVCGGKNSCRKYSGLTP 699  
674 QPKVVDGTPCSPPDSSTSVCCQCVKAGCDRIIDSKKFKDKGCGGNGSTCKKISGVS 733  
700 TNYGNDIVTIPAGATNIDVKORSHPGVONDGNATLAKTADGOYLNLGNLAISAIBODIL 759  
734 AKPGYHDIITFTGATNIEVQKORNGSRNNGSFLAKADGTIILNDDTLSTLEBDM 793  
760 VKGTLKXSGSIATLERLQSFRLPEPLTVOLLTVPEVPPKXKYTFEVPNDVDSMOS 819  
794 YKGVLYRSGSSAALERISFSPKPEPLTIOVLTV-GNALRPKIKYVYFV----- 842  
820 SKERATNIIQPLHMQVLTGDMWSCSSTCGAGMORRTVECRDPSGGAATCNALKPDE 879  
843 KKKESFPAI-PTGSA-WVIEBWGSCSKCELGMRRLVECRDINGOPASCAKEVAPAS 900  
880 AKPCESQCP 889  
901 TRPCADHPCP 910  
RESULT 10  
US-10-115-286-2  
Sequence 2, Application US/10115286  
Publication No. US2003016065A1  
GENERAL INFORMATION:  
APPLICANT: Jonak, Zdenka  
Trull, Stephen  
Formwald, James  
Teirle, Jonathan  
Hasting, Gregg  
TITLE OF INVENTION: No. US2003016065A1 Integrin Ligand ITGL-TSP  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/115,286  
FILING DATE: 04-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/845,496

FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul P  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846189  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 967 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-115-286-2  
Query Match 48.0%; Score 2328.5; DB 14; Length 967;  
Best Local Similarity 49.7%; Pred. No. 3,4e-176;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;  
11 LPTLLLLLLPLPARGAPAPAGGASSELVVP--TRLPGSAGELALHLSAFKGFVLR 68  
33 VPTLLLLAALLAVS-DALGRPE--EDELVVPELERAPGH-GTTRRLHAFDQDLLE 88  
69 LAPDSSFLAPEFKIERLGGSGRATGQ-----ERGLRGCFSGGTVNGEPELSLAVALSICRGL 123  
89 LRPDSSFLAPGFTLQNV--GRKSGSEPLPEPTDLAHCFYSGTVNGDPSSAALSLICEGV 145  
124 SGSFLIDGEEFTTIOPOGAGGSLAOPRILOKMPAPARPLP-----IGPEMEV---ET 172  
146 RGAFLYLGSAVFIQPPAAS-----ERLATAAGKRPAPLOFHLRRRQGVGTGCV 200  
173 GEGORQERGDHODESSEESCEEAEGASBP-----PLGATLS-RTKSFVSEARFV 222  
201 VDDEPRPTGALETDEDEDETEGEDEGPQWSPODPALQVQGTGTGSIKKKFFVSHRYV 260  
223 ETLVADASMAAFYGDILQNHILITLMSVARIYKPSIKNSINLMVYKLYVEDEKMGSE 282  
261 ETMLVADQSMAEHFGSLGKHYLLITLPSVAARYKPSINSVSLVVKLVHDEKMGSE 320  
283 VSDNGGLTLRNFQNMORREFNOPSDRHEHYDTAILTRONFCGOBGLCDTLGADIGTIC 342  
321 VTSNAALTNRFCWQKXHPSPDRDAEHYDIALFTRODLGSGQ--TCDTLGMADVGTVC 379  
343 DPNKSCSVIEDBGLQAAHTLAHEIGHVLSMPHDSKPCTRLFGPMKHHVMAPLFVHLNQ 402  
380 DPNKSCSVIEDDGLQAAFTTAHEIGHVFNMPHDDAKOCASLNGNODSHMMAKMLNLDH 439  
403 TLPMSPCSAMVLTLLDGGHDCILDAFGALPLPTGLPGRMALYQDDQCRQIFGDFR 462  
440 SQPWSPCSAMVLTSLDNGHGBCLMDKPNPQLPGLPG--TSYANRQCCGTFGBDCK 497  
463 HCPNTSADVCAQIMC-HTDGAEPILCHTKNGSLPMWADGTPCGPGLHCSESSCLPEEVER 521  
498 HCPD--AASCTSLMCTGTSGGVAVLCQTKH--FPWADGTSCEGKMCINKCVAKTDKRA 553  
522 PKPVVDGGMAMPWGMGESSRTCCGGVQFNSHRECKDEPONGGRYCLGRARAKYOSCHTEEC 581  
554 FDTFFHSGMWGMWGMGDCSRTCCGGVQYTMRECDNVPKNGGKICBGRVRYRSCNLEDC 613  
582 PP-DGKSFREOQCEKXNAVNTDM-DGNLLQWPKYAGVSPRDRCKLFCRARGSEPKYF 639  
614 PDNNGKTFREOCEAHNBFSSKASFGSGPAVEMIPKAVGSPKRCKLICQAKIGYFVL 673  
640 EAKYIDGTLGPEPTLAICVRGQCVKAGCDHVVDSPKLDKCGVCGGKNSCRKYSGLTP 699  
674 QPKVVDGTPCSPPDSSTSVCCQCVKAGCDRIIDSKKFKDKGCGGNGSTCKKISGVS 733  
700 TNYGNDIVTIPAGATNIDVKORSHPGVONDGNATLAKTADGOYLNLGNLAISAIBODIL 759

```
Db 734 AKPGYHDIITITPGATNIEVKQRNQRGSRNNGSFLAIKAADGTIILNGDYTLSTLEEDIM 793
Qy 760 VKGTLIKSGSIATLERISFRPLPEPLTVOLLTVPEVPPPKYKTFEFPVNDVDSMOS 819
Db 794 YKGVNLRYSGSSALERIRSPFLKEPLTIQVLTIV-GNALRFXIKITYEV----- 842
Qy 820 SKERATNIIQPLIHAQWVLGDWSESSSTCGAGMQRRTVECRDPSGOASATCNKALPED 879
Db 843 KKKKESFNAI-PTFSA-WVIEEMGECSCSLGMQRRLVECRDINGQDPASCAKEVPAS 900
Qy 880 AKPCEBQICP 889
Db 901 TRPCADHPCP 910

RESULT 11
US-09-373-658-125
; Sequence 125, Application US/09373658
; Publication No. US20030092900A1
; GENERAL INFORMATION:
; APPLICANT: Iruela-Arispe, Luisa
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Jonak, Zdenka L.
; APPLICANT: Trulli, Stephen H.
; APPLICANT: Fromwald, James A.
; APPLICANT: Terrett, Jonathan A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.1070006
; CURRENT APPLICATION NUMBER: US/09/373,658
; NUMBER OF SEQ ID NOS: 125
; CURRENT FILING DATE: 1999-08-13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 125
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-373-658-125

Query Match 48.0%; Score 2328.5; DB 10; Length 968;
Best Local Similarity 49.7%; Pred. No. 3.4e-176;
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

Qy 11 LPEFLILLPLLRGAPAPAPAGQASSELVVP--TRLPGSAGELALHLSAFGKGFVLR 68
Db 34 VPLILLALAAVLAVS-DALGRPSE--EDELIVPELEBARGH-GTTRLRIHAFOQDLLE 89
Qy 69 LAPDDSLAPFEKIERLGGSGRATGQ---ERGLRGCFPSGTVNGEPESLAASLCKGL 123
Db 90 LRDDSSFLAPGFLLQNV--GRKSGSETPLPETDLAHCFYSGTVNGDPSAAALSLCKGV 146
Qy 124 SSGFLLDGEFTIOPGAGGSLAQPHRLQRMGPAGARPLP-----RGPMEVY---ET 172
Db 147 RGAFFYLLGAYFAFQPLPAAS-----ERLATAAPGEXKPPAPLQFHLIRNRQDVGTCGV 201
Qy 173 GSGQROERGDHQRDESEESOEERAGASBP-----PPLGATS-RTRKFSVBARFV 222
Db 202 VDDEPRPTGAETREDEDEGTGEDEDEGPQMSPOPALQGVGQGTGSIIRKRFVSSHVYV 261
Qy 223 ETLVLVADASAAATGADLQNHITLMSVAARIYKHSINOSINIMVVKYLIEDKMPRE 282
Db 262 ETLVLVADASAAATGADLQNHITLMSVAARIYKHSINOSINIMVVKYLIEDKMPRE 321
Qy 283 VSDNGGLTLNFCNMQRRFNQPSDRAPHEHYDTAIIITRONFCGQEGLCDTLGVADIGTIC 342
Db 322 VTISNALLTLNFCNMQRRFNQPSDRAPHEHYDTAIIITRONFCGQEGLCDTLGVADIGTIC 380
Qy 343 DPKKSGSVLEDEGLQAHTLAHLGHTLSNPHDSSPCRTLPGMKKHVAPLPHVNLQ 402
Db 381 DPKKSGSVLEDEGLQAHTLAHLGHTLSNPHDSSPCRTLPGMKKHVAPLPHVNLQ 440
Qy 403 TLWSPGSAWYLTLELLDGGHGDCLDAPGAALPLPTGLPGRNALYOLDOQCQOIFQDPFR 462
```

```
Db 441 SQPMSEPCSAWYMTSLFLNGHGECLMDKPNPIQLFEDLPG--TSYDANKQOCFTGEDSK 498
Qy 463 HCPNTSAQDVCAQLWC-HTDGAELPLCHTRNGSLPFWADGTPCGPHLCSEBGLPEEYER 521
Db 499 HCPD--NASTGSLTCTGTSGLVLCQTKH--FPAADGTSCEBGMCTNGKCVNTRDKH 554
Qy 522 PKPVVDGMAWPGMGECSRTGCGVQFSHCECDPEFQNGRXYCLGRARAKQSHTEC 581
Db 555 FDTPLFGSGVMGMPMGDCSRCTGCGVQYMECDPVKNQGGKCYEGKVRKRSCTLEDC 614
Qy 582 PP-DGKSFREOQCEKNANVTDM--DGNLLQWVPKXAGVSPEDRCKLFCRAGRSEFYFV 639
Db 615 PDNNGKTFREBOCEANERSKASPSGSPRAVENIPIKXAGVSPEDRCKLFCQAKGIGYFVL 674
Qy 640 EAKVIDGTLGPEETALCYRGQCVTAQCDHVVDSFRKLDKCGVCGGKSGSKVSGSLTP 699
Db 675 CPKRVVDGTPCSPDSTGCVQCVKAGCDRIIDSKKEDKCGVCGANSTCKKISGSYTS 734
Qy 700 TWYGVNDIYITPAGATNIDVKORSHPGVQNDGNVYALATKADQYLLNGIALAISAEODIL 759
Db 735 AKPGYHDIITITPGATNIEVKQRNQRGSRNNGSFLAIKAADGTIILNGDYTLSTLEEDIM 794
Qy 760 VKGTLIKSGSIATLERISFRPLPEPLTVOLLTVPEVPPPKYKTFEFPVNDVDSMOS 819
Db 795 YKGVNLRYSGSSALERIRSPFLKEPLTIQVLTIV-GNALRFXIKITYEV----- 843
Qy 820 SKERATNIIQPLIHAQWVLGDWSESSSTCGAGMQRRTVECRDPSGOASATCNKALPED 879
Db 844 KKKKESFNAI-PTFSA-WVIEEMGECSCSLGMQRRLVECRDINGQDPASCAKEVPAS 901
Qy 880 AKPCEBQICP 889
Db 902 TRPCADHPCP 911

RESULT 12
US-09-741-151-4
; Sequence 4, Application US/09741151
; Publication No. US20020086400A1
; GENERAL INFORMATION:
; APPLICANT: ZHU, Shiaooping et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NOCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO01.005
; CURRENT APPLICATION NUMBER: US/09/741,151
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Human
US-09-741-151-4

Query Match 48.0%; Score 2328; DB 12; Length 931;
Best Local Similarity 49.8%; Pred. No. 3.6e-176;
Matches 451; Conservative 143; Mismatches 250; Indels 62; Gaps 21;

Qy 15 LILLILLPLLRGAPAPAPAGQASSELVVP--TRLPGSAGELALHLSAFGKGFVLR 72
Db 1 LILLILLPLLRGAPAPAPAGQASSELVVP--TRLPGSAGELALHLSAFGKGFVLR 72
Qy 73 DGLAPFEKIERLGGSGRATGQ---ERGLRGCFPSGTVNGEPESLAASLCKGLSGSF 127
Db 58 SSFLAPFEKIERLGGSGRATGQ---ERGLRGCFPSGTVNGEPESLAASLCKGLSGSF 114
Qy 128 LLDGEFTIOPGAGGSLAQPHRLQRMGPAGARPLP-----RGPMEVY---ETGEGQ 176
Db 115 YLDEAFYIQLPPLAS-----ERLATAAPGEXKPPAPLQFHLIRNRQDVGTCGVYDDE 169
Qy 177 ROERGHQDESEESOEERAGASBP-----PPLGATS-RTRKFSVBARFVETLL 226
```

```

Db      170 PRETKAETDEDEGEDEGEDEGAQMSFODFALQGVQPTGTGSIKKRFVSSHRYETML 229
Qy      227 VAAASMAAFYADLQNHILTLMSVAARIYKGPSIKNSINIMVYKVLIVEDEKMWPEVSDN 286
Db      230 VAAQSAAEFFSGSLKHYLLTLFSVAARLYHPSIRNSVSLVYVKILVIHDEQGPATYN 289
Qy      287 GGLTFNFCNMQRRFNQPSDRHPEHYDTALLTRQNFQGEGLCDTLGVADIGTICDPNK 346
Db      290 AALLTRNFCNMQRRFNQPSDRHPEHYDTALLTRQNFQGEGLCDTLGVADIGTICDPNK 348
Qy      347 SCSTIEBEGLOAATTAHGLHVLMSPHDSCPTLFGPMGKHMAPLFVHNLQTLPM 406
Db      349 SCSTIEBEGLOAATTAHGLHVLMSPHDSCPTLFGPMGKHMAPLFVHNLQTLPM 408
Qy      407 SPCSAMYTELLDGHDCCLIDAFGAALPLPTGLPGRMALYQDDQCRQIFGDFHHCN 466
Db      409 SPCGAMYTELLDGHDCCLIDAFGAALPLPTGLPGRMALYQDDQCRQIFGDFHHCN 466
Qy      467 TSAQVCAQLWC-HTDGAEPULCHTKNSLPMADGTPCGPHLCSGSCLEPHEVERPKV 525
Db      467 -AASTCSTLMCTGTSGVLVCCQTKH--FPMADGTSGBGKWCINGCNVAKTDRKHPDTP 522
Qy      526 VDGWAMFMPWGECSRTCCGGVOPSHRECKDPEPONGGRYCLGRRAKYQSCHEECP-D 584
Db      523 FHGWMGMWMPGDCSRTCCGGVOPSHRECKDPEPONGGRYCLGRRAKYQSCHEECP-D 582
Qy      585 GKSFREOQCEKYNANYTDM-DGNLQWPKYAGVSPRDKLFCPARGRSEFKFEAKY 643
Db      583 GKTFREOQCEKYNANYTDM-DGNLQWPKYAGVSPRDKLFCPARGRSEFKFEAKY 642
Qy      644 IDGLQCPBETALIVRGQCVKAGCDHVDSBRKLDKCGVCGKNSCRKYSGLTPTNYG 703
Db      643 VDGPPCPDSTSVCCVQCVKAGCDRIIDSKKFKDKGVCGNGSTCKKISSGVTSAKRG 702
Qy      704 YNDIVTTPAGATNIDVQKSHPGVONGDNIALKTAQGYLLNGNLAISIBODILVKG 763
Db      703 YNDIVTTPAGATNIDVQKSHPGVONGDNIALKTAQGYLLNGNLAISIBODILVKG 762
Qy      764 ILKSGSATIERLQSRPLPEPTVOLLTVPGVFPKXKTFEPVNDVDFSMOSK 823
Db      763 VLKSGSATIERLQSRPLPEPTVOLLTVPGVFPKXKTFEPVNDVDFSMOSK 823
Qy      824 ATTNIQPLHAQWVLGWSCESSCTCGAGQRRYECRDPGQASATCNKALXPEDAKPC 883
Db      812 ESFNAI-PTFSA-WVIEWGECSKSCJELGWRRLVECRDINGQASCAEAVDPASTR 869
Qy      884 ESQUCP 889
Db      870 ADHPCP 875

RESULT 13
US-09-989-687-126
; Sequence 126, Application US/09989687
; Publication No. US20040002449A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 126
; LENGTH: 967
; TYPE: PRT
; ORGANISM: ITGL-TSP
US-09-989-687-126

Query Match      47.8%; Score 2320.5; DB 11; Length 967;
Best Local Similarity 49.5%; Pred. No. 1.5e-175;
Matches 450; Conservative 147; Mismatches 250; Indels 63; Gaps 22;

```

```

Qy      11 LPFLILLILLPLARGAPARPAAGQASSELVVP--TRLPSSAGELALHLSAFGKGYLR 68
Db      33 VPTILLALLAALAAVS-DALGRPSE--EDEELIVLELRFVPH-GTTRLRHAFFQOQDLD 88
Qy      69 LAPDSEFLAEPFKIERLGSGRATG-----ERLGRGFFSGTNGEESLAASLRCGL 123
Db      89 VPDSEFLAEPFKIERLGSGRATG-----ERLGRGFFSGTNGEESLAASLRCGL 145
Qy      124 SGSPFLDGEFTTPOGAGSIAQPHRLQRMGPAARLP-----RGPEWY---ET 172
Db      146 RGAFYLGEAFYIQLPAAAS-----ERLATAPEKPEPAPLQFHLRRNRQGDVGATCGV 200
Qy      173 GEGQRQERQDQDESEESQEEBAEASERP-----PPLATLS-RTGRFVSEARFV 222
Db      201 VDDEPRFTKAETDEDEGEDEGEDEGAQMSFODFALQGVQPTGTGSIKKRFVSSHRY 260
Qy      223 ETLVADASMAAFYADLQNHILTLMSVAARIYKGPSIKNSINIMVYKVLIVEDEKWP 282
Db      261 ETLVADASMAAFYADLQNHILTLMSVAARIYKGPSIKNSINIMVYKVLIVEDEKWP 320
Qy      283 VSDNGGLTLKMFQWRRFNQPSDRHPEHYDTALLTRQNFQGEGLCDTLGVADIGTIC 342
Db      321 VTSNAALTLRNFQWRRFNQPSDRHPEHYDTALLTRQNFQGEGLCDTLGVADIGTIC 379
Qy      343 DPNKSCSYIEBEGLOAATTAHGLHVLMSPHDSCPTLFGPMGKHMAPLFVHNLQ 402
Db      380 DPNKSCSYIEBEGLOAATTAHGLHVLMSPHDSCPTLFGPMGKHMAPLFVHNLQ 439
Qy      403 TLPWSPCSAMYTELLDGHDCCLIDAFGAALPLPTGLPGRMALYQDDQCRQIFGDFH 462
Db      440 SPMWSPCSAMYTELLDGHDCCLIDAFGAALPLPTGLPGRMALYQDDQCRQIFGDFH 497
Qy      463 HCPNMSADQVCAQLWC-HTDGAEPULCHTKNSLPMADGTPCGPHLCSGSCLEPHE 521
Db      498 HCPNMSADQVCAQLWC-HTDGAEPULCHTKNSLPMADGTPCGPHLCSGSCLEPHE 553
Qy      522 PKPVVDGAMFMPWGECSRTCCGGVOPSHRECKDPEPONGGRYCLGRRAKYQSCHE 581
Db      554 PKPVVDGAMFMPWGECSRTCCGGVOPSHRECKDPEPONGGRYCLGRRAKYQSCHE 613
Qy      582 PP-DGKSFREOQCEKYNANYTDM-DGNLQWPKYAGVSPRDKLFCPARGRSEFKFE 639
Db      614 PP-DGKSFREOQCEKYNANYTDM-DGNLQWPKYAGVSPRDKLFCPARGRSEFKFE 673
Qy      640 EAKVIDGTLQCPBETALIVRGQCVKAGCDHVDSBRKLDKCGVCGKNSCRKYSGLT 699
Db      674 EAKVIDGTLQCPBETALIVRGQCVKAGCDHVDSBRKLDKCGVCGKNSCRKYSGLT 733
Qy      700 TNYGNDIVTTPAGATNIDVQKSHPGVONGDNIALKTAQGYLLNGNLAISIBODIL 759
Db      734 TNYGNDIVTTPAGATNIDVQKSHPGVONGDNIALKTAQGYLLNGNLAISIBODIL 793
Qy      760 VKGTIKYSGSATIERLQSRPLPEPTVOLLTVPGVFPKXKTFEPVNDVDFSMOS 819
Db      794 VKGTIKYSGSATIERLQSRPLPEPTVOLLTVPGVFPKXKTFEPVNDVDFSMOS 842
Qy      820 SKERATTNIQPLHAQWVLGWSCESSCTCGAGQRRYECRDPGQASATCNKALXPED 879
Db      843 SKERATTNIQPLHAQWVLGWSCESSCTCGAGQRRYECRDPGQASATCNKALXPED 900
Qy      880 AKPCESQUCP 889
Db      901 TRPCADHPCP 910

RESULT 14
US-09-802-582-8
; Sequence 8, Application US/09802582
; Publication No. US20020086354A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.

```

```

; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; FILE REFERENCE: 0734-323001
; CURRENT APPLICATION NUMBER: US 09/802,582
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-802-582-8

```

```

Query Match      46.9%; Score 2274.5; DB 12; Length 481;
Best Local Similarity 82.1%; Pred. No. 2.5e-172;
Matches 403; Conservative 35; Mismatches 40; Indels 13; Gaps 2;

```

```

QY 401 NOTLPMSPCSAMYLTELDDGHDCLDPAFGALPLPTGLPGRMALYOLDQGCQIFGPD 460
DB 3 NKTLPWSPCSAYLYTELDDGHDCLD-----DGHSTLYELDDQGCQIFGPD 50
QY 461 FRHCNTSAQDYCAQLWC-HTDGAFLCHTNGSLPMADGTPCGPGLHCSBGSLPEEY 519
DB 51 FRHCNTSVEIDCVLWCRHRSDDEPCHTNGASLMDGTPCGPGLHCSBGSLPEEY 110
QY 520 EEPKPVVDGMAFMGPWGECSRTCCGGVQFSHRECKDEPONGRYCLGRAXYQSCHE 579
DB 111 ENPKAVVDGDMGPMGPWGCSTCCGGIQFSNRECDNAPONGRFLCGERVKTQSCKE 170
QY 580 ECPDGSFRBOQCEKNAVNTTMDGNLQWVPKYAGVSPDRCKLFCRARGSEFKVF 639
DB 171 ECPDGSFRBOQCEKNAVNTTMDGNLQWVPKYAGVSPDRCKLFCRARGSEFKVF 230
QY 640 EAKVIDGTLGPEETALICVRGQCVKAGCDHYVDSPRKLDKCGVCGGKNSCRKXGSLTP 699
DB 231 ETKVIDGTLGPEETALICVRGQCVKAGCDHYVDSPRKLDKCGVCGGKNSCRKXGSLTP 290
QY 700 TNYGNDIVITIPAGATNIDVKORSHPGVQNDGNVYALKTADQOYLINGNLASAIEDIL 759
DB 291 PSYGNDIVITIPAGATNIDVKORSHPGVQNDGSYALKTANGQYLINGNLASAIEDIL 350
QY 760 VKGTLIKYSGSIATLERLOSFRPLPEPLTVOLLTVGSEVFPKXYTFVFNVDVFSMOS 819
DB 351 MKGTLIKYSGSMATLERLOSFOALPEPLTVOLLTVGSEVFPKXYTFVFNVDVFSMOS 410
QY 820 SKERATNTIIOPLLAQWYLGDMSECSSTCGAGMORRYTECRDPGQASATCNKALKPED 879
DB 411 SKERASTNTIIOPLLAQWYLGDMSECSSTCGAGMORRYTECRDPGQASATCNKALKPED 470
QY 880 AKPCESQQLCPL 890
DB 471 AKPCGSQQLCPL 481

```

```

RESULT 15
US-10-105-929-8
; Sequence 8, Application US/10105929

```

```

; Publication No. US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US 10/105,929
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus rattus
; US-10-105-929-8

```

```

Query Match      46.9%; Score 2274.5; DB 13; Length 481;
Best Local Similarity 82.1%; Pred. No. 2.5e-172;
Matches 403; Conservative 35; Mismatches 40; Indels 13; Gaps 2;

```

```

QY 401 NOTLPMSPCSAMYLTELDDGHDCLDPAFGALPLPTGLPGRMALYOLDQGCQIFGPD 460
DB 3 NKTLPWSPCSAYLYTELDDGHDCLD-----DGHSTLYELDDQGCQIFGPD 50
QY 461 FRHCNTSAQDYCAQLWC-HTDGAFLCHTNGSLPMADGTPCGPGLHCSBGSLPEEY 519
DB 51 FRHCNTSVEIDCVLWCRHRSDDEPCHTNGASLMDGTPCGPGLHCSBGSLPEEY 110
QY 520 EEPKPVVDGMAFMGPWGECSRTCCGGVQFSHRECKDEPONGRYCLGRAXYQSCHE 579
DB 111 ENPKAVVDGDMGPMGPWGCSTCCGGIQFSNRECDNAPONGRFLCGERVKTQSCKE 170
QY 580 ECPDGSFRBOQCEKNAVNTTMDGNLQWVPKYAGVSPDRCKLFCRARGSEFKVF 639
DB 171 ECPDGSFRBOQCEKNAVNTTMDGNLQWVPKYAGVSPDRCKLFCRARGSEFKVF 230
QY 640 EAKVIDGTLGPEETALICVRGQCVKAGCDHYVDSPRKLDKCGVCGGKNSCRKXGSLTP 699
DB 231 ETKVIDGTLGPEETALICVRGQCVKAGCDHYVDSPRKLDKCGVCGGKNSCRKXGSLTP 290
QY 700 TNYGNDIVITIPAGATNIDVKORSHPGVQNDGNVYALKTADQOYLINGNLASAIEDIL 759
DB 291 PSYGNDIVITIPAGATNIDVKORSHPGVQNDGSYALKTANGQYLINGNLASAIEDIL 350
QY 760 VKGTLIKYSGSIATLERLOSFRPLPEPLTVOLLTVGSEVFPKXYTFVFNVDVFSMOS 819
DB 351 MKGTLIKYSGSMATLERLOSFOALPEPLTVOLLTVGSEVFPKXYTFVFNVDVFSMOS 410
QY 820 SKERATNTIIOPLLAQWYLGDMSECSSTCGAGMORRYTECRDPGQASATCNKALKPED 879
DB 411 SKERASTNTIIOPLLAQWYLGDMSECSSTCGAGMORRYTECRDPGQASATCNKALKPED 470
QY 880 AKPCESQQLCPL 890
DB 471 AKPCGSQQLCPL 481

```

```

Search completed: May 7, 2004, 11:57:19
Job time : 58 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:56:21 ; Search time 26 Seconds  
(without alignments)  
3292.710 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890

Sequence: 1 MFPAAPAPRWLPULLLLLL.....CNRALKPEDAKPCESQLCLP 890

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	1.5	2165	2	T21371
2	12	1.3	205	2	S37804
3	12	1.3	837	2	T00355
4	12	1.3	951	2	T00017
5	11	1.2	263	2	SS7346
6	11	1.2	655	1	A46688
7	10	1.1	231	2	A49265
8	10	1.1	331	2	T15587
9	10	1.1	480	2	A31589
10	10	1.1	491	2	UC6197
11	10	1.1	492	2	A44399
12	10	1.1	550	2	T47158
13	10	1.1	1015	2	UC5263
14	10	1.1	1015	2	UC5062
15	10	1.1	1277	2	T30532
16	10	1.1	112	2	A60111
17	9	1.0	133	2	SS7038
18	9	1.0	134	2	A40959
19	9	1.0	144	2	T18867
20	9	1.0	171	2	A15502
21	9	1.0	182	2	UC7236
22	9	1.0	185	2	A49957
23	9	1.0	189	2	UC7262
24	9	1.0	197	2	SS1372
25	9	1.0	206	2	A55412
26	9	1.0	244	2	S18948
27	9	1.0	285	2	UC5639
28	9	1.0	348	2	G02297
29	9	1.0	368	2	AEO086

30	9	1.0	381	2	S35940	class I histocompa
31	9	1.0	382	2	B84527	hypothetical prote
32	9	1.0	384	2	S25771	gaal protein - mou
33	9	1.0	402	1	BOHUS	sex steroid-bindin
34	9	1.0	409	2	H70303	ubiquinol-cytochro
35	9	1.0	429	2	T17215	hypothetical prote
36	9	1.0	452	2	F71327	probable signal re
37	9	1.0	481	2	T22406	hypothetical prote
38	9	1.0	500	2	S65522	cartilage matrix p
39	9	1.0	557	2	S12359	nicotinic acetylch
40	9	1.0	602	2	S69198	prostoglandin G/H
41	9	1.0	602	2	S39782	cyclooxigenase 1 -
42	9	1.0	626	1	NBRHUA	platelet glycoprot
43	9	1.0	871	2	S47518	cadherin - African
44	9	1.0	941	2	TYMVMD	protein-tyrosine k
45	9	1.0	964	2	T15746	hypothetical prote

#### ALIGNMENTS

RESULT 1  
T21371  
hypothetical protein F25H8.3 - Caenorhabditis elegans  
C/Spectrum: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T21371, T24896  
R/Gadastly, S.  
submitted to the EMBL Data Library, February 1996  
A/Reference number: Z19413  
A/Accession: T21371  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-2165 <M12>  
A/Cross-references: EMBL:Z69360; PIDD:CA93287.1; GSPDB:GN00022; CESP:F25H8.3  
R/Gadastly, S.  
submitted to the EMBL Data Library, February 1996  
A/Reference number: Z19949  
A/Accession: T24896  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-2165 <M12>  
A/Cross-references: EMBL:Z69361; PIDD:CA93288.1; GSPDB:GN00022; CESP:F25H8.3  
A/Experimental source: clone T13H10  
C/Genetics:  
A/Map position: 4  
A/Insertions: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 8

Query Match  
Best Local Similarity 100.0%; Pred. No. 0.00083;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 MGECSTRTCGGQVQ 548  
DB 612 MGECSTRTCGGQVQ 624

RESULT 2  
S37804  
hypothetical protein YKL47C - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein YKL601  
C/Spectrum: Saccharomyces cerevisiae  
C/Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Apr-2002  
R/Vandenbol, M.; Boile, P.; Dion, C.; Porteballe, D.; Hilger, F.  
submitted to the EMBL Data Library, September 1993  
A/Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAF4 1  
A/Accession: S37804  
A/Molecule type: DNA  
A/Residues: 1-205 <VAN>

A/Cross-references: EMBL:Z26877; NID:9407482; PIDN:CAA81507.1; PID:9407501  
 A/Experimental source: Strain S288C  
 R/Vanderhol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.  
 Submitted to the Protein Sequence Database, March 1994  
 A/Reference number: S37976  
 A/Accession: S37977  
 A/Molecule type: DNA  
 A/Residues: 1-205 <VA2>  
 A/Cross-references: EMBL:Z28146; NID:9486250; PIDN:CAA81987.1; PID:9486251; MIPS:YKL147C  
 A/Experimental source: Strain S288C  
 R/Vanderhol, M.; Bolle, P.A.; Dion, C.; Portetelle, D.; Hilger, F.  
 Yeast 10, 35-40, 1994  
 A/Title: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LBP4 loci of  
 A/Reference number: S44563  
 A/Accession: S44581  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-205 <VA3>  
 A/Cross-references: EMBL:Z26877; NID:9407482; PIDN:CAA81507.1; PID:9407501  
 A/Experimental source: Strain S288C  
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1993  
 C/Genetics:  
 A/Cross-references: SGD:S0001630  
 A/Map position: 11L  
 C/Superfamily: Saccharomyces hypothetical protein YKL147C

Query Match 1.3%; Score 12; DB 2; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 0.0012;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 FLLLLLLPL 24  
 DB 73 FLLLLLLPL 84

RESULT 3  
 T00355  
 Hypothetical protein KIAA0688 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 21-Jul-2000  
 C/Accession: T00355  
 R/ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.  
 DNA Res. 5, 169-176, 1998  
 A/Title: Prediction of the coding sequences of unidentified human genes. X. The complete  
 A/Reference number: Z14142; MUID:98403880; PMID:9734811  
 A/Accession: T00355  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-837 <ISH>  
 A/Cross-references: EMBL:AB014588; NID:93327189; PIDN:BA03163.1; PID:93327190  
 A/Experimental source: brain  
 A/Genetics:  
 A/Gene: KIAA0688  
 C/Superfamily: thrombospondin type 1 repeat homology <THR3>  
 F:519-575/Domain: thrombospondin type 1 repeat homology

Query Match 1.3%; Score 12; DB 2; Length 837;  
 Best Local Similarity 100.0%; Pred. No. 0.0038;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTGGGVQFS 550  
 DB 532 CSRTGGGVQFS 543

RESULT 4  
 T00017  
 gene ADAMTS-1 protein - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 21-Jul-2000  
 C/Accession: T00017  
 R/Kono, K.; Iizasa, H.; Ohno, S.; Matsushima, K.  
 Genomics 46, 466-471, 1997

A/Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gen  
 A/Reference number: Z14055; MUID:98110583; PMID:9441751  
 A/Accession: T00017  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-951 <KUN>  
 A/Cross-references: EMBL:AB001735; NID:92809056; PIDN:BA04501.1; PID:92809057  
 A/Experimental source: strain 129SVJ  
 A/Genetics:  
 A/Gene: ADAMTS-1  
 A/Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2  
 C/Superfamily: thrombospondin type 1 repeat homology  
 F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 1.3%; Score 12; DB 2; Length 951;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 DIVTIPAGATNI 717  
 DB 724 DIVTIPAGATNI 735

RESULT 5  
 S57346  
 Interleukin 15 receptor precursor - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 05-Nov-1999  
 C/Accession: S57346  
 R/Giri, U.G.; Kumaki, S.; Ahdieh, M.; Friend, D.C.; Loomis, A.; Shanebeck, K.; Dubose,  
 EMO J. 14, 3654-3663, 1995  
 A/Title: Identification and cloning of a novel IL-15 binding protein that is structural  
 A/Reference number: S57346; MUID:95369237; PMID:7641685  
 A/Accession: S57346  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-263 <GIR>  
 A/Cross-references: EMBL:U22339; NID:9551105; PIDN:AA052240.1; PID:9551106  
 C/Superfamily: complement factor H repeat homology  
 C/Keywords: cytokine receptor  
 F:36-96/Domain: complement factor H repeat homology <FH2>

Query Match 1.2%; Score 11; DB 2; Length 263;  
 Best Local Similarity 100.0%; Pred. No. 0.015;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLPL 24  
 DB 18 LLLLLLLPL 28

RESULT 6  
 A46688  
 hepatocyte growth factor activator (BC 3.4.21.-) precursor [validated] - human  
 C/Species: Homo sapiens (man)  
 C/Date: 21-Sep-1993 #sequence\_revision 25-Aug-1995 #text\_change 08-Dec-2000  
 C/Accession: A46688  
 R/Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.  
 J. Biol. Chem. 268, 10024-10028, 1993  
 A/Title: Molecular cloning and sequence analysis of the cDNA for a human serine protea  
 d coagulation factor XII.  
 A/Reference number: A46688; MUID:93252879; PMID:7683665  
 A/Accession: A46688  
 A/Molecule type: mRNA  
 A/Residues: 1-655 <MTY>  
 A/Cross-references: DBJ:D14012; NID:9219680; PIDN:BA03113.1; PID:9219681  
 A/Experimental source: liver (mRNA); serum (protein)  
 A/Note: sequence extracted from NCBI backbone (NCBI:131227, NCBI:131228)  
 C/Genetics:  
 A/Gene: GDB:HGFAC; HGPA: HGPA  
 A/Cross-references: GDB:954514  
 A/Map position: 4p16-4p16

C/Function: activates hepatocyte growth factor by specific proteolytic cleavage  
 A/Pathway: tissue repair and regeneration  
 C/Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;  
 C/Keywords: glycoprotein; hydrolase; Kringle; liver; plasma; serine proteinase  
 F/1-34/Domain: signal sequence status predicted <Sig>  
 F/108-148/Domain: fibronectin type II repeat homology <LF2>  
 F/164-197/Domain: EGF homology <EG1>  
 F/202-237/Domain: fibronectin type I repeat homology <LF1>  
 F/245-278/Domain: EGF homology <EG2>  
 F/286-367/Domain: Kringle homology <KRG>  
 F/373-407/Domain: hepatocyte growth factor activator light chain #status experimental  
 F/408-655/Domain: hepatocyte growth factor activator heavy chain #status experimental  
 F/408-661/Domain: trypsin homology <TRY>  
 F/440-48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F/164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-36  
 F/447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 1.2%; Score 11; DB 1; Length 655;  
 Best Local Similarity 100.0%; Pred. No. 0.031;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 PFLLLLLLLL 22  
 DB 18 PFLLLLLLLL 28

## RESULT 7

A49265  
 C/Species: Mus musculus (house mouse)  
 C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 08-Oct-1999  
 C/Accession: A49265; 149347; 149346; S43290  
 R/Lyman, S.D.; James, L.; Vanden Bos, T.; de Vries, P.; Brasel, K.; Gliniak, B.; Holling  
 D.; Williams, D.E.; Beckmann, M.P.  
 A/Title: Molecular cloning of a ligand for the flt3/flk-2 tyrosine kinase receptor: a p  
 A/Reference number: A49265; MUID:94084791; PMID:7505204  
 A/Accession: A49265  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-1231 <LTM>  
 A/Cross-references: GB:U23636; NID:g439441; PID:AAA93436.1; PID:g439442  
 R/Lyman, S.D.; Stocking, K.; Davison, B.; Fletcher, F.; Johnson, L.; Escobar, S.  
 Oncogene 11, 1165-1172, 1995  
 A/Title: Structural analysis of human and murine flt3 ligand genomic loci.  
 A/Reference number: 139075; MUID:96032581; PMID:7566977  
 A/Accession: 149347  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-163; G, 165, 'HAG' <RES>  
 A/Cross-references: EMBL:U29875; NID:g1072039; PID:AAA90952.1; PID:g1072041  
 A/Accession: 149346  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-197, 'L', 198-231 <RES>  
 A/Cross-references: EMBL:U29875; NID:g1072039; PID:AAA90951.1; PID:g1072040  
 R/Hanum, C.; Culpepper, J.; Campbell, D.; McClanahan, T.; Zurawski, S.; Bazan, J.F.; Ka  
 felt, A.; Menich, M.; Kelnier, G.; Namiyawa, R.; Kemick, D.; Roncetto, M.G.; Zlotnik, A  
 Nature 368, 643-648, 1994  
 A/Title: Ligand for FLT3/Flk2 receptor tyrosine kinase regulates growth of haematopoiet  
 A/Reference number: S43290; MUID:94195428; PMID:8145851  
 A/Accession: S43290  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-197, 'L', 198-231 <HAN>  
 A/Experimental source: clone T110  
 A/Note: the sequence from Fig. 2c is inconsistent with that from Fig. 2a in having 4-Val  
 C/Genetics: 11/3; 49/3; 67/3; 115/3; 164/1; 224/3  
 C/Keywords: transmembrane protein

Query Match

1.1%; Score 10; DB 2; Length 231;

Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LLLLLLLPL 24  
 DB 190 LLLLLLLPL 199

## RESULT 8

hypothetical protein C24A3.3 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C/Accession: T15587  
 R/Favell, T.  
 submitted to the EMBL Data Library, November 1995  
 A/Description: The sequence of C. elegans cosmid C24A3.  
 A/Reference number: Z18373

A/Accession: T15587  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-331 <FAV>  
 A/Cross-references: EMBL:U40424; NID:g1065542; PID:g1065546; PID:AAA81458.1; CESP:C24A  
 C/Genetics: 27/1; 73/2; 136/1; 193/2  
 A/Introns: 27/1; 73/2; 136/1; 193/2

Query Match 1.1%; Score 10; DB 2; Length 331;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 LLLLLLLPL 23  
 DB 320 LLLLLLLPL 329

## RESULT 9

A31589  
 C/Species: Homo sapiens (man)  
 C/Date: 11-May-1989 #sequence\_revision 11-May-1989 #text\_change 05-May-2000  
 C/Accession: A31589  
 R/Galljart, N.J.; Gillemaus, N.; Harris, A.; van der Horst, G.T.J.; Verheijen, F.W.; Gal  
 Cell 54, 755-764, 1988  
 A/Title: Expression of cDNA encoding the human "protective protein" associated with lys  
 A/Reference number: A31589; MUID:88311078; PMID:3136930  
 A/Accession: A31589  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-480 <GAL>  
 A/Cross-references: GB:M22960; GB:J03159; GB:M18453; NID:g190282; PID:AAA36476.1; PID:  
 C/Superfamily: serine carboxypeptidase  
 C/Keywords: hydrolase; serine carboxypeptidase  
 F/176,400,457/Active site: Ser, Asp, His #status predicted

Query Match 1.1%; Score 10; DB 2; Length 480;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 PLLLLLLLL 22  
 DB 10 PLLLLLLLL 19

## RESULT 10

JC6197  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 17-Mar-1999  
 C/Accession: JC6197  
 R/Okada, A.; Saez, S.; Mismul, Y.; Basset, P.  
 Gene 185, 187-193, 1997



A/Title: Rat stromelysin 3: cDNA cloning from healing skin wound, activation by furin at  
 A/Reference number: J06197; MUID:97208872; PMID:9055814  
 A/Contents: Skin wounds  
 A/Accession: J06197  
 A/Molecule type: mRNA  
 A/Residues: 1-491 <OKA>  
 A/Cross-references: GB:U66034  
 C/Comment: This protein is a member of the matrix metalloproteinase family.  
 C/Superfamily: Interstitial collagenase; hemopexin repeat homology; matrix metalloprote  
 C/Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F:52-261/Domain: matrix metalloproteinase homology <MMP>  
 F:291-483/Domain: hemopexin repeat homology <PXN>  
 F:184,218,222,228/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:218,222,228/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:219/Active site: Glu #status predicted

Query Match 1.1%; Score 10; DB 2; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLP 23  
 Db 21 LLLLLLLLP 30

RESULT 11  
 A44399  
 stromelysin 3 (EC 3.4.24.-) - mouse  
 N/Alternate names: matrix metalloproteinase 11 (MMP11)  
 C/Species: Mus musculus (house mouse)  
 C/Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Jul-1998  
 C/Accession: A44399  
 R/Refby: O.; Wolf, C.; Linacher, J.M.; Hulin, P.; Wendling, C.; Lemeur, M.; Basset, F  
 J. Cell Biol. 119, 997-1002, 1992  
 A/Title: The breast cancer-associated stromelysin-3 gene is expressed during mouse mamma  
 A/Reference number: A44399; MUID:93054930; PMID:142845  
 A/Accession: A44399  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: nucleic acid  
 A/Residues: 1-493 <LEF>  
 A/Note: sequence extracted from NCBI backbone (NCBIP:117216)  
 C/Superfamily: Interstitial collagenase; hemopexin repeat homology; matrix metalloprotei  
 C/Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F:52-262/Domain: matrix metalloproteinase homology <MMP>  
 F:292-484/Domain: hemopexin repeat homology <PXN>  
 F:184,219,223,229/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:219,223,229/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:220/Active site: Glu #status predicted

Query Match 1.1%; Score 10; DB 2; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLP 23  
 Db 21 LLLLLLLLP 30

RESULT 12  
 T47158  
 hypothetical protein DKFp762C110.1 - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C/Accession: T47158  
 R/Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.  
 submitted to the Protein Sequence Database, March 2000  
 A/Reference number: Z24379  
 A/Accession: T47158  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-550 <AAA>  
 A/Cross-references: EMBL:AL162080  
 A/Experimental source: adult melanoma (W60 cell line); clone DKFp762C110

C/Genetics:  
 A/Note: DKFp762C110.1

Query Match 1.1%; Score 10; DB 2; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.28;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGQ 548  
 Db 154 CSRTCGGQ 163

RESULT 13  
 JC5263  
 transmembrane tyrosine phosphatase-like protein, ICAR - human  
 C/Species: Homo sapiens (man)  
 C/Date: 25-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 20-Jun-2000  
 C/Accession: JC5263  
 R/Smith, P.D.; Barker, K.T.; Wang, J.; Lu, Y.J.; Shipley, J.; Crompton, M.R.  
 Biochem. Biophys. Res. Commun. 229, 402-411, 1996  
 A/Title: ICAR, a novel member of a new family of transmembrane, tyrosine phosphatase-1  
 A/Reference number: JC5263; MUID:97127415; PMID:8954911  
 A/Accession: JC5263  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-1015 <SMI>  
 A/Cross-references: GB:Y08569; NID:G1644377; PID:CAA69880.1; PID:G1644378  
 C/Comment: This protein has an intracellular protein tyrosine phosphatase like protein.  
 C/Superfamily: protein-tyrosine-phosphatase, receptor type N; protein-tyrosine-phosphat  
 C/Keywords: phosphoprotein  
 F:770-994/Domain: protein-tyrosine-phosphatase homology <PTP>  
 F:945/Active site: Cys (phosphocysteine intermediate) #status predicted  
 F:951/Binding site: substrate phosphate (Arg) #status predicted

Query Match 1.1%; Score 10; DB 2; Length 1015;  
 Best Local Similarity 100.0%; Pred. No. 0.47;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLP 23  
 Db 7 LLLLLLLLP 16

RESULT 14  
 JC5062  
 phogrin precursor - human  
 N/Contents: protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type  
 C/Species: Homo sapiens (man)  
 C/Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 21-Jun-2002  
 C/Accession: JC5062; JC5263; T46903  
 R/Kawasaki, E.; Hutton, J.C.; Bisenbath, G.S.  
 Biochem. Biophys. Res. Commun. 227, 440-447, 1996  
 A/Title: Molecular cloning and characterization of the human transmembrane protein tyr  
 A/Reference number: JC5062; MUID:97032784; PMID:8878534  
 A/Accession: JC5062  
 A/Contents: islet  
 A/Molecule type: mRNA  
 A/Residues: 1-1015 <RAW>  
 A/Cross-references: GB:U66702; NID:G1620663; PID:AA50742.1; PID:G1620664  
 R/Smith, P.D.; Barker, K.T.; Wang, J.; Lu, Y.J.; Shipley, J.; Crompton, M.R.  
 Biochem. Biophys. Res. Commun. 229, 402-411, 1996  
 A/Title: ICAR, a novel member of a new family of transmembrane, tyrosine phosphatase-1  
 A/Reference number: JC5263; MUID:97127415; PMID:8954911  
 A/Accession: JC5263  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-207, 'S', 209-246, 'G', 248-322, 'R', 324, 'N', 326-1015 <SMI>  
 A/Cross-references: GB:Y08569; NID:G1644377; PID:CAA69880.1; PID:G1644378  
 R/Jansorge, W.; Winkler, U.; Mewes, H.W.; Weil, B.; Wiemann, S.  
 submitted to the Protein Sequence Database, February 2000  
 A/Reference number: Z24134  
 A/Accession: T46903  
 A/Status: preliminary

A:Molecule type: mRNA  
 A:Residues: 'DA',714,771-1015 <AAA>  
 A:Cross-references: EMBL:AL157451  
 A:Experimental source: adult amygdala; clone DKFZp761A0712  
 C:Comment: This protein has an intracellular protein tyrosine phosphatase like protein.  
 C:Genetics:  
 A>Note: DKFZp761A0712.1  
 C:Superfamily: protein-tyrosine-phosphatase, receptor type N; protein-tyrosine-phosphatase  
 C:Keywords: phosphoprotein; phosphoric monoester hydrolase; transmembrane protein; tyrosine  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-1015/Product: phogrin #status predicted <MAT>  
 F:615-639/Domain: transmembrane #status predicted <TM>  
 F:770-994/Domain: protein-tyrosine-phosphatase homology <PTP2>  
 F:945/Active site: Cys (phosphocysteine intermediate) #status predicted  
 F:951/Binding site: substrate phosphate (Arg) #status predicted

Query Match 1.1%; Score 10; DB 2; Length 1015;  
 Best Local Similarity 100.0%; Pred. No. 0.47; Mismatches 0; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLP 23  
 |||||  
 DB 7 LLLLLLLLP 16

## RESULT 15

T30532  
 neural cell adhesion molecule LI homolog - Fugu rubripes

C:Species: Fugu rubripes

C>Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 01-Mar-2002

C:Accession: T30532

R:Ripoldi Tunicliffe, G.R.; Plutzer, M.; Nyakatura, G.; Elgar, G.S.; Brenner, S.; Rosen  
 Submitted to the EMBL Data Library, September 1997

A:Description: Analysis of the genomic loci of Fugu rubripes homologs of the human disc

A:Reference number: Z20848

A:Accession: T30532

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1277 <RIB>

A:Cross-references: EMBL:AF026198, NID:G3098263, PID:G3098264, PIDN:AACT5580.1

C:Genetics:

A:Introns: 42/1; 47/1; 81/2; 149/1; 190/1; 247/1; 285/2; 347/1; 391/1; 440/1; 477/2; 531/2

A>Note: LI-CAM

C:Superfamily: neural cell adhesion molecule LI; fibronectin type III repeat homology; I

Query Match 1.1%; Score 10; DB 2; Length 1277;  
 Best Local Similarity 100.0%; Pred. No. 0.57;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLLLLLLPLA 25  
 |||||  
 DB 19 LLLLLLLPLA 28

Search completed: May 7, 2004, 12:00:48  
 Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 12:00:12 ; Search time 54 Seconds  
(without alignments)  
4574.711 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890  
Sequence: 1 MFPAAPAPRWLPFLILL.....CNKALXPDAPKPCESQLCPL 890

Scoring table: CLIGO  
Gapop 60.0, Gapext 60.0

Searched: 1140673 seqs, 27566755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	890	100.0	890	10	US-09-373-658-4
2	890	100.0	890	11	US-09-989-687-4
3	710	79.8	924	12	US-10-425-114-39107
4	263	29.6	364	9	US-09-764-903-57
5	217	24.4	245	9	US-09-918-171A-11
6	35	3.9	905	9	US-09-918-171A-9
7	30	3.4	481	12	US-09-802-583-8
8	30	3.4	481	13	US-10-105-929-8
9	30	3.4	481	14	US-10-363-227-8
10	13	1.5	2150	9	US-09-321-87B-2
11	13	1.5	2165	9	US-09-800-729-155
12	12	1.3	438	9	US-09-963-791-22
13	12	1.3	438	12	US-10-419-276-22
14	12	1.3	518	9	US-09-803-589-10
15	12	1.3	551	12	US-09-802-582-16

16	12	1.3	551	13	US-10-105-929-16	Sequence 16, Appl
17	12	1.3	551	14	US-10-365-227-16	Sequence 16, Appl
18	12	1.3	575	12	US-10-358-283-12	Sequence 12, Appl
19	12	1.3	589	9	US-09-963-791-12	Sequence 12, Appl
20	12	1.3	589	12	US-10-419-276-12	Sequence 12, Appl
21	12	1.3	727	9	US-09-445-023A-12	Sequence 12, Appl
22	12	1.3	727	14	US-10-097-597-12	Sequence 12, Appl
23	12	1.3	727	14	US-10-097-580-12	Sequence 12, Appl
24	12	1.3	757	9	US-09-963-791-24	Sequence 24, Appl
25	12	1.3	757	12	US-10-419-276-24	Sequence 24, Appl
26	12	1.3	837	10	US-09-946-374-317	Sequence 317, App
27	12	1.3	837	12	US-10-206-915-352	Sequence 352, App
28	12	1.3	837	12	US-10-199-670-352	Sequence 352, App
29	12	1.3	837	12	US-10-201-858-352	Sequence 352, App
30	12	1.3	837	12	US-10-205-890-352	Sequence 352, App
31	12	1.3	837	12	US-10-208-024-352	Sequence 352, App
32	12	1.3	837	12	US-10-201-853-352	Sequence 352, App
33	12	1.3	837	12	US-10-358-283-11	Sequence 11, Appl
34	12	1.3	837	12	US-10-174-581-352	Sequence 352, App
35	12	1.3	837	12	US-10-176-483-352	Sequence 352, App
36	12	1.3	837	12	US-10-176-749-352	Sequence 352, App
37	12	1.3	837	12	US-10-176-914-352	Sequence 352, App
38	12	1.3	837	12	US-10-176-915-352	Sequence 352, App
39	12	1.3	837	12	US-10-106-485A-317	Sequence 317, App
40	12	1.3	837	12	US-10-013-907A-317	Sequence 317, App
41	12	1.3	837	12	US-10-015-499A-317	Sequence 317, App
42	12	1.3	837	12	US-10-176-484-352	Sequence 352, App
43	12	1.3	837	12	US-10-180-550-352	Sequence 352, App
44	12	1.3	837	12	US-10-183-014-352	Sequence 352, App
45	12	1.3	837	12	US-10-187-738-352	Sequence 352, App

#### ALIGNMENTS

RESULT 1  
US-09-373-658-4  
Sequence 4, Application US/09373658  
Publication No. US20030092900A1  
GENERAL INFORMATION:  
APPLICANT: Ituela-Artepe, Luisa  
APPLICANT: Haestings, Gregg A.  
APPLICANT: Ruben, Steven M.  
APPLICANT: Jonak, Zdenek L.  
APPLICANT: Trull, Stephen H.  
APPLICANT: Fromwald, James A.  
TITLE OF INVENTION: Meth and Meth2 Polynucleotides and Polypeptides  
FILE REFERENCE: 1488.107006  
CURRENT APPLICATION NUMBER: US/09/373,658  
CURRENT FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 125  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 890  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-373-658-4

Query Match 100.0%; Score 890; DB 10; Length 890;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPAAPAPRWLPFLILLPLARGAPAPPAAGGASSELVPTRLPGSAGELATHISA 60  
DB 1 MFPAAPAPRWLPFLILLPLARGAPAPPAAGGASSELVPTRLPGSAGELATHISA 60  
QY 61 FKGKGVRLADDSFLAPDFKIERLGGSGRATGGERLGRGTFSGTYNGPESLAAYSLC 120  
DB 61 FKGKGVRLADDSFLAPDFKIERLGGSGRATGGERLGRGTFSGTYNGPESLAAYSLC 120  
QY 121 RGLSGFLDDEERTIOPQAGAGSLAOPHRICRWGPAGARLPFGPWEVETGEGORER 180  
DB 121 RGLSGFLDDEERTIOPQAGAGSLAOPHRICRWGPAGARLPFGPWEVETGEGORER 180

```

Db      121 RGLSGSFLDGESEFTIOPQAGAGSLAOPHRLQRMWGPAGARELPGRPEMEVETGEGORER 180
Qy      181 GDHODESESESEAEAGASBPPLGATSRTRKFEVSEARFVETLLVADASMAAFYADL 240
Db      181 GDHODESESESEAEAGASBPPLGATSRTRKFEVSEARFVETLLVADASMAAFYADL 240
Qy      241 QNHILTLMSVARIYKHPISIKNSINLMVYKYLIVDEKMGPEVSDNGLLTLRNFQNMQR 300
Db      241 QNHILTLMSVARIYKHPISIKNSINLMVYKYLIVDEKMGPEVSDNGLLTLRNFQNMQR 300
Qy      301 FNQPSDRPHEHYDTAILLTRONFCQEGELCDTLGVADIGTICDPNKSQSVIIEDEGLQAAH 360
Db      301 FNQPSDRPHEHYDTAILLTRONFCQEGELCDTLGVADIGTICDPNKSQSVIIEDEGLQAAH 360
Qy      361 TLAEHLGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLPMSPCSAWYLTLLDG 420
Db      361 TLAEHLGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLPMSPCSAWYLTLLDG 420
Qy      421 GHGDCILDPAGALPLPTGLPGRMALYOLDQCCROI FGPDRHCNPTSAODVCAQLMCHT 480
Db      421 GHGDCILDPAGALPLPTGLPGRMALYOLDQCCROI FGPDRHCNPTSAODVCAQLMCHT 480
Qy      481 DGAEPILCHTKNGSLPMAOGTFCGPHLCSSESCLEPEEVEERPKPVVDGMAIPMGWEGCS 540
Db      481 DGAEPILCHTKNGSLPMAOGTFCGPHLCSSESCLEPEEVEERPKPVVDGMAIPMGWEGCS 540
Qy      541 RTGGGVGFVSHRECKDPEPONGRCLGRRAKYOSCHTEBCPPDGKSFREOQCEKXNAYN 600
Db      541 RTGGGVGFVSHRECKDPEPONGRCLGRRAKYOSCHTEBCPPDGKSFREOQCEKXNAYN 600
Qy      601 YTDMDGNLQWVPKXAGVSPDRCLFCRARSEPFKFEAKVIDGTLGCEPTELLAVRG 660
Db      601 YTDMDGNLQWVPKXAGVSPDRCLFCRARSEPFKFEAKVIDGTLGCEPTELLAVRG 660
Qy      661 QCVKAGCDHVVDSPKLDKCGVCGGKNSCRKVSGLTPTNYGYNDIYTIIPAGATNIDVK 720
Db      661 QCVKAGCDHVVDSPKLDKCGVCGGKNSCRKVSGLTPTNYGYNDIYTIIPAGATNIDVK 720
Qy      721 QRSHPGVNDGNYLALKTADQYLLNGMLAISAIODILVKTILKXGSIATLERIOSF 780
Db      721 QRSHPGVNDGNYLALKTADQYLLNGMLAISAIODILVKTILKXGSIATLERIOSF 780
Qy      781 RPLPEPLTVOLLTVGEVEFPKXKTYTFVNDVDFSMOSKEXRATNIIQPLLAHQWVLG 840
Db      781 RPLPEPLTVOLLTVGEVEFPKXKTYTFVNDVDFSMOSKEXRATNIIQPLLAHQWVLG 840
Qy      841 DMSBGSSTCGAGMORRTVECRDPGQASATCNKALKPEDAKPCESQLCPL 890
Db      841 DMSBGSSTCGAGMORRTVECRDPGQASATCNKALKPEDAKPCESQLCPL 890

```

## RESULT 2

```

US-09-989-687-4
; Sequence 4, Application US/09989667
; Publication No. US20040002449A1
; GENERAL INFORMATION:
; APPLICANT: Hassting, Gregg A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 890
; ORGANISM: Homo sapiens
US-09-989-687-4

```

```

Query Match      100.0%; Score 890; DB 11; Length 890;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MFPAPAPAPWLPFLLLLLLLLLLLLLLPLARGAPAPAPAGQASLWVPTRLPGSAGELALHLSA 60
Db      1 MFPAPAPAPWLPFLLLLLLLLLLLLLLPLARGAPAPAPAGQASLWVPTRLPGSAGELALHLSA 60
Qy      61 FGKGYLRLAPDSDTLAPFKIEERLGGSGRATGGRGIRGCFSGTUNVGEPSLAAVSLC 120
Db      61 FGKGYLRLAPDSDTLAPFKIEERLGGSGRATGGRGIRGCFSGTUNVGEPSLAAVSLC 120
Qy      61 FGKGYLRLAPDSDTLAPFKIEERLGGSGRATGGRGIRGCFSGTUNVGEPSLAAVSLC 120
Db      61 FGKGYLRLAPDSDTLAPFKIEERLGGSGRATGGRGIRGCFSGTUNVGEPSLAAVSLC 120
Qy      121 RGLSGSFLDGESEFTIOPQAGAGSLAOPHRLQRMWGPAGARELPGRPEMEVETGEGORER 180
Db      121 RGLSGSFLDGESEFTIOPQAGAGSLAOPHRLQRMWGPAGARELPGRPEMEVETGEGORER 180
Qy      181 GDHODESESESEAEAGASBPPLGATSRTRKFEVSEARFVETLLVADASMAAFYADL 240
Db      181 GDHODESESESEAEAGASBPPLGATSRTRKFEVSEARFVETLLVADASMAAFYADL 240
Qy      241 QNHILTLMSVARIYKHPISIKNSINLMVYKYLIVDEKMGPEVSDNGLLTLRNFQNMQR 300
Db      241 QNHILTLMSVARIYKHPISIKNSINLMVYKYLIVDEKMGPEVSDNGLLTLRNFQNMQR 300
Qy      301 FNQPSDRPHEHYDTAILLTRONFCQEGELCDTLGVADIGTICDPNKSQSVIIEDEGLQAAH 360
Db      301 FNQPSDRPHEHYDTAILLTRONFCQEGELCDTLGVADIGTICDPNKSQSVIIEDEGLQAAH 360
Qy      361 TLAEHLGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLPMSPCSAWYLTLLDG 420
Db      361 TLAEHLGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLPMSPCSAWYLTLLDG 420
Qy      421 GHGDCILDPAGALPLPTGLPGRMALYOLDQCCROI FGPDRHCNPTSAODVCAQLMCHT 480
Db      421 GHGDCILDPAGALPLPTGLPGRMALYOLDQCCROI FGPDRHCNPTSAODVCAQLMCHT 480
Qy      481 DGAEPILCHTKNGSLPMAOGTFCGPHLCSSESCLEPEEVEERPKPVVDGMAIPMGWEGCS 540
Db      481 DGAEPILCHTKNGSLPMAOGTFCGPHLCSSESCLEPEEVEERPKPVVDGMAIPMGWEGCS 540
Qy      541 RTGGGVGFVSHRECKDPEPONGRCLGRRAKYOSCHTEBCPPDGKSFREOQCEKXNAYN 600
Db      541 RTGGGVGFVSHRECKDPEPONGRCLGRRAKYOSCHTEBCPPDGKSFREOQCEKXNAYN 600
Qy      601 YTDMDGNLQWVPKXAGVSPDRCLFCRARSEPFKFEAKVIDGTLGCEPTELLAVRG 660
Db      601 YTDMDGNLQWVPKXAGVSPDRCLFCRARSEPFKFEAKVIDGTLGCEPTELLAVRG 660
Qy      661 QCVKAGCDHVVDSPKLDKCGVCGGKNSCRKVSGLTPTNYGYNDIYTIIPAGATNIDVK 720
Db      661 QCVKAGCDHVVDSPKLDKCGVCGGKNSCRKVSGLTPTNYGYNDIYTIIPAGATNIDVK 720
Qy      721 QRSHPGVNDGNYLALKTADQYLLNGMLAISAIODILVKTILKXGSIATLERIOSF 780
Db      721 QRSHPGVNDGNYLALKTADQYLLNGMLAISAIODILVKTILKXGSIATLERIOSF 780
Qy      781 RPLPEPLTVOLLTVGEVEFPKXKTYTFVNDVDFSMOSKEXRATNIIQPLLAHQWVLG 840
Db      781 RPLPEPLTVOLLTVGEVEFPKXKTYTFVNDVDFSMOSKEXRATNIIQPLLAHQWVLG 840
Qy      841 DMSBGSSTCGAGMORRTVECRDPGQASATCNKALKPEDAKPCESQLCPL 890
Db      841 DMSBGSSTCGAGMORRTVECRDPGQASATCNKALKPEDAKPCESQLCPL 890

```

## RESULT 3

```

US-10-425-114-39107
; Sequence 39107, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E
; APPLICANT: Cao, Yongwei

```

;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
;; FILE REFERENCE: 38-21(53313) B  
;; CURRENT APPLICATION NUMBER: US/10/425,114  
;; CURRENT FILING DATE: 2003-04-28  
;; NUMBER OF SEQ ID NOS: 73128  
;; SEQ ID NO 39107  
;; LENGTH: 924  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATUURE:  
;; OTHER INFORMATION: Clone ID: LIB4654-025-F6\_F11.pcp  
US-10-425-114-39107

Query Match 79.8%; Score 710; DB 12; Length 924;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 810; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 80 FRIERLGGSGRATGGERGLRGCFPSGTNGEPESLAVALSICRLGSGFLDGEFTIQP 139
DB 114 FRIERLGGSGRATGGERGLRGCFPSGTNGEPESLAVALSICRLGSGFLDGEFTIQP 173
QY 140 GAGGSIAQPRRLORWGPAGARPLPRGPEWEVENGEORORGDHDSSEESQEEAAGA 199
DB 174 GAGGSIAQPRRLORWGPAGARPLPRGPEWEVENGEORORGDHDSSEESQEEAAGA 233
QY 200 SEPPPLGATSRTRKFVSEARFVETLLVADASMAAFYGADLQNHILLMSVARIYKHS 259
DB 234 SEPPPLGATSRTRKFVSEARFVETLLVADASMAAFYGADLQNHILLMSVARIYKHS 293
QY 260 IKKSNIMVYKULIVBEKKGPEVSDNGSLTFNFCQMGRRFPQSDRDRPHYDTAILLT 319
DB 294 IKKSNIMVYKULIVBEKKGPEVSDNGSLTFNFCQMGRRFPQSDRDRPHYDTAILLT 353
QY 320 RQNFCEGEGLCDTLGVADIGTICDPNKSQVIEDGLQAAHTLAHGLVLSMFHDSKP 379
DB 354 RQNFCEGEGLCDTLGVADIGTICDPNKSQVIEDGLQAAHTLAHGLVLSMFHDSKP 413
QY 380 CTRLFGMGKHVYAPLFVHLNOTLPMSPCSAMYLTEBLDGHGDCLLDAPGALPLPTG 439
DB 414 CTRLFGMGKHVYAPLFVHLNOTLPMSPCSAMYLTEBLDGHGDCLLDAPGALPLPTG 473
QY 440 LPRGMALYQDQOCROQIFGPDFRHCPTNSAQDVCAQJMCHTDGAEPPLCKTNGSLPMADG 499
DB 474 LPRGMALYQDQOCROQIFGPDFRHCPTNSAQDVCAQJMCHTDGAEPPLCKTNGSLPMADG 533
QY 500 TPCGPGHLCSEGSCLPREEVERKPPVVDGGMAPMGWEGCSRTCGGVOFSHRECKDPP 559
DB 534 TPCGPGHLCSEGSCLPREEVERKPPVVDGGMAPMGWEGCSRTCGGVOFSHRECKDPP 593
QY 560 QNGRXYCLGRARAYQSCHEECPPDQKSPREOCCEKXNANNTYDMGNLLQWPKKAGS 619
DB 594 QNGRXYCLGRARAYQSCHEECPPDQKSPREOCCEKXNANNTYDMGNLLQWPKKAGS 653
QY 620 PRDRCKLFCRARGSEKVFVAVYDGLTGPETLAIYVRGQCVKAGCDHVVDSPKLDK 679
DB 654 PRDRCKLFCRARGSEKVFVAVYDGLTGPETLAIYVRGQCVKAGCDHVVDSPKLDK 713
QY 680 CGVCGGAGNSCRKXSGSLPTNYGNDIVTTPAGANTIDYKQSHGVDGNDGYLALKTA 739
DB 714 CGVCGGAGNSCRKXSGSLPTNYGNDIVTTPAGANTIDYKQSHGVDGNDGYLALKTA 773
QY 740 DGOYLNGNLAIASIEDILVKGITLKYSIATLEBSFPLPEPLTVQLLTVEGEV 799
DB 774 DGOYLNGNLAIASIEDILVKGITLKYSIATLEBSFPLPEPLTVQLLTVEGEV 833
QY 800 PPKYKTFYFVNDVDFSMQSKERATNTIIOPLIHAQWVLGDMSECSGTGAGMQRRTVE 859
DB 834 PPKYKTFYFVNDVDFSMQSKERATNTIIOPLIHAQWVLGDMSECSGTGAGMQRRTVE 893
QY 860 CRDPSGQASATCNALKPEDAKPCESQICPL 890
DB 894 CRDPSGQASATCNALKPEDAKPCESQICPL 924

```

RESULT 4  
US-09-764-903-57  
Sequence 57, Application US/09764903  
Patent No. US2002030674A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: P7128  
CURRENT APPLICATION NUMBER: US/09/764,903  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 57  
LENGTH: 364  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-903-57

Query Match 29.6%; Score 263; DB 9; Length 364;  
Best Local Similarity 100.0%; Pred. No. 9,9e-232; Indels 0; Gaps 0;  
Matches 263; Conservative 0; Mismatches 0;

```

QY 530 WAPWPMGECSCRTCGGVOFSHRECKDPEFONGRYCLGRARAYQSCHEECPPDGKSF 589
DB 4 WAPWPMGECSCRTCGGVOFSHRECKDPEFONGRYCLGRARAYQSCHEECPPDGKSF 63
QY 590 EOCCEKXNANNTYDMGNLLQWPKKAGVSPRDSCKLFCRARGSEKVFVAVYDGLTLC 649
DB 64 EOCCEKXNANNTYDMGNLLQWPKKAGVSPRDSCKLFCRARGSEKVFVAVYDGLTLC 123
QY 650 GPETLAIYVRGQCVKAGCDHVVDSPKLDKCGVCGKNSCRKXSGSLPTNYGNDIVT 709
DB 124 GPETLAIYVRGQCVKAGCDHVVDSPKLDKCGVCGKNSCRKXSGSLPTNYGNDIVT 183
QY 710 IPAGANTIDYKQSHGVDGNDGYLALKTAQGYLLNGNLAIASIEDILVKGITLKYSG 769
DB 184 IPAGANTIDYKQSHGVDGNDGYLALKTAQGYLLNGNLAIASIEDILVKGITLKYSG 243
QY 770 SIATLERLQSFRLPEPLTVQLL 792
DB 244 SIATLERLQSFRLPEPLTVQLL 266

```

RESULT 5  
US-09-918-171A-11  
Sequence 11, Application US/09918171A  
Patent No. US20020110894A1  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurekainen, Tiina L.  
APPLICANT: Hitchaba, Satoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/04193  
CURRENT APPLICATION NUMBER: US/09/918,171A  
CURRENT FILING DATE: 2001-07-30  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 11  
LENGTH: 245  
TYPE: PRT  
ORGANISM: Homo sapiens ADAMTS-8  
US-09-918-171A-11

Query Match 24.4%; Score 217; DB 9; Length 245;  
Best Local Similarity 100.0%; Pred. No. 8.5e-190;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 196 AEGASEPPPLGATSRTRKFVSEARFVETLLVADASMAAFYGADLQNHILLMSVARIY 255

```

DB 2 AEGASEPPPLGATSRTRKFEVSEARFVETLLVADSMAPFADLQNNITLMSVARIY 61
QY 256 KHPISIKSINILNVAVVLLVEDEKMGPEVSDNGGLTRNFCMQRFPNPSDRHPEHYDTA 315
DB 62 KHPISIKSINILNVAVVLLVEDEKMGPEVSDNGGLTRNFCMQRFPNPSDRHPEHYDTA 121
QY 316 ILITRONFCGGGGLCDTIGVADIGTICDPNKSCTVIEBEGLOAAATLAHELGHVLSMFD 375
DB 122 ILITRONFCGGGGLCDTIGVADIGTICDPNKSCTVIEBEGLOAAATLAHELGHVLSMFD 181
QY 376 DSKPCTRLFGPMGKHVNAFLFVHNLQTLPMSPCSAM 412
DB 182 DSKPCTRLFGPMGKHVNAFLFVHNLQTLPMSPCSAM 218

RESULT 6
US-09-918-171A-9
; Sequence 9, Application US/09918171A
; Patent No. US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apce, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/04193
; CURRENT APPLICATION NUMBER: US/09/918,171A
; PRIORITY FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIORITY FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 905
; TYPE: PRT
; ORGANISM: Mus musculus ADAMTS-8
US-09-918-171A-9

Query Match
Best Local Similarity 100.0%; Pred. No. 9,48-23;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 GVSPPDRCKLFCRARGSEFFKVFPAKVDTGTCGP 651
DB 632 GVSPPDRCKLFCRARGSEFFKVFPAKVDTGTCGP 666

RESULT 7
US-09-802-582-8
; Sequence 8, Application US/09802582
; Publication No. US20020086354A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-323001
; CURRENT APPLICATION NUMBER: US/09/802,582
; PRIORITY FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIORITY FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01

```

```

; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-802-582-8

Query Match
Best Local Similarity 100.0%; Pred. No. 2e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 YGYNDIVTIPAGATNIDVQRSHPGVQNDG 731
DB 293 YGYNDIVTIPAGATNIDVQRSHPGVQNDG 322

RESULT 8
US-10-105-929-8
; Sequence 8, Application US/10105929
; Publication No. US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/10/105,929
; PRIORITY FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus rattus
US-10-105-929-8

Query Match
Best Local Similarity 100.0%; Pred. No. 2e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 YGYNDIVTIPAGATNIDVQRSHPGVQNDG 731
DB 293 YGYNDIVTIPAGATNIDVQRSHPGVQNDG 322

RESULT 9
US-10-365-227-8
; Sequence 8, Application US/10365227
; Publication No. US20030143632A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-323001
; CURRENT APPLICATION NUMBER: US/10/365,227
; PRIORITY FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: US/09/802,582
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04

```

PRIOR APPLICATION NUMBER: US 09/130,491  
PRIOR FILING DATE: 1998-08-06  
PRIOR APPLICATION NUMBER: US 60/054,966  
PRIOR FILING DATE: 1997-08-06  
PRIOR APPLICATION NUMBER: US 60/058,108  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: US 09/388,280  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: US 09/388,279  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 481  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-10-365-227-8

Query Match 3.4%; Score 30; DB 14; Length 481;  
Best Local Similarity 100.0%; Pred. No. 2e-18;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 YGYNDIVTIPAGATNIVKQSHPGVQNDG 731  
DB 293 YGYNDIVTIPAGATNIVKQSHPGVQNDG 322

RESULT 10  
US-09-321-987B-2  
Sequence 2, Application US/09321987B  
Patent No. US20020102210A1  
GENERAL INFORMATION:  
APPLICANT: Blueloch, Robert H  
TITLE OF INVENTION: Agent and Method for Modulating Cell Migration  
FILE REFERENCE: 960296.95386  
CURRENT FILING DATE: 1999-05-28  
PRIOR APPLICATION NUMBER: US/09/321,987B  
PRIOR FILING DATE: 1998-05-29  
PRIOR FILING DATE: 1998-05-29  
PRIOR APPLICATION NUMBER: 60/129,023  
PRIOR FILING DATE: 1999-04-13  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2150  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-321-987B-2

Query Match 1.5%; Score 13; DB 9; Length 2150;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGVQ 548  
DB 587 WGECSRTCGGVQ 609

RESULT 11  
US-09-800-729-155  
Sequence 155, Application US/09800729  
Patent No. US20020068319A1  
GENERAL INFORMATION:  
APPLICANT: Ni et al.  
TITLE OF INVENTION: 32 Human secreted proteins  
FILE REFERENCE: P2044P1  
CURRENT FILING DATE: US/09/800,729  
CURRENT FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: PCT/US00/26013  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 60/155,709  
PRIOR FILING DATE: 1999-09-24

NUMBER OF SEQ ID NOS: 217  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 155  
LENGTH: 2165  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-800-729-155

Query Match 1.5%; Score 13; DB 9; Length 2165;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGVQ 548  
DB 612 WGECSRTCGGVQ 624

RESULT 12  
US-09-963-791-22  
Sequence 22, Application US/09963791  
Patent No. US20020120113A1  
GENERAL INFORMATION:  
APPLICANT: Donoho, Gregory  
APPLICANT: Turner, C. Alexander Jr.  
APPLICANT: Friedrich, Glenn  
APPLICANT: Scoville, John  
APPLICANT: Zambrowicz, Brian  
TITLE OF INVENTION: No. US20020120113A1 Human Proteases and Polynucleotides Encodi  
FILE REFERENCE: LEX-0105-USA  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: US 60/169,769  
PRIOR FILING DATE: 1999-12-09  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 438  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-963-791-22

Query Match 1.3%; Score 12; DB 9; Length 438;  
Best Local Similarity 100.0%; Pred. No. 0.054;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGV 547  
DB 416 WGECSRTCGGV 427

RESULT 13  
US-10-419-276-22  
Sequence 22, Application US/10419276  
Publication No. US20030225258A1  
GENERAL INFORMATION:  
APPLICANT: Donoho, Gregory  
APPLICANT: Turner, C. Alexander Jr.  
APPLICANT: Friedrich, Glenn  
APPLICANT: Scoville, John  
APPLICANT: Zambrowicz, Brian  
TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same  
FILE REFERENCE: LEX-0105-USA  
CURRENT FILING DATE: US/10/419,276  
CURRENT FILING DATE: 2003-04-17  
PRIOR APPLICATION NUMBER: US/09/963,791  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: US 60/169,769  
PRIOR FILING DATE: 1999-12-09  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22



LENGTH: 438  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-419-276-22

Query Match 1.3%; Score 12; DB 12; Length 438;  
 Best Local Similarity 100.0%; Pred. No. 0.054;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGV 547  
 DB 416 WGECSRTCGGV 427

RESULT 14  
 US-09-803-589-10  
 Sequence 10, Application US/09803589  
 Patent No. US20020112251A1  
 GENERAL INFORMATION:  
 APPLICANT: McCarthy, Sean A.  
 APPLICANT: Holtzman, Douglas A.  
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
 TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER  
 TITLE OF INVENTION: USES  
 FILE REFERENCE: 07334-325001  
 CURRENT APPLICATION NUMBER: US/09/803,589  
 CURRENT FILING DATE: 2001-03-09  
 PRIOR APPLICATION NUMBER: US 09/128,709  
 PRIOR FILING DATE: 1998-08-04  
 PRIOR APPLICATION NUMBER: US 60/054,645  
 PRIOR FILING DATE: 1997-08-04  
 PRIOR APPLICATION NUMBER: US 09/130,491  
 PRIOR FILING DATE: 1998-08-06  
 PRIOR APPLICATION NUMBER: US 60/054,966  
 PRIOR FILING DATE: 1997-08-06  
 PRIOR APPLICATION NUMBER: US 60/058,108  
 PRIOR FILING DATE: 1997-09-05  
 PRIOR APPLICATION NUMBER: US 09/388,280  
 PRIOR FILING DATE: 1999-09-01  
 PRIOR APPLICATION NUMBER: US 09/388,279  
 PRIOR FILING DATE: 1999-09-01  
 NUMBER OF SEQ ID NOS: 14  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 10  
 LENGTH: 518  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-803-589-10

Query Match 1.3%; Score 12; DB 9; Length 518;  
 Best Local Similarity 100.0%; Pred. No. 0.063;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 DIVTIPAGATNI 717  
 DB 291 DIVTIPAGATNI 302

RESULT 15  
 US-09-802-582-16  
 Sequence 16, Application US/09802582  
 Publication No. US20020086354A1  
 GENERAL INFORMATION:  
 APPLICANT: McCarthy, Sean A.  
 APPLICANT: Holtzman, Douglas A.  
 APPLICANT: Goodheart, Andrew D.J.  
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
 TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER  
 TITLE OF INVENTION: USES  
 FILE REFERENCE: 07334-323001  
 CURRENT APPLICATION NUMBER: US/09/802,582  
 CURRENT FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: US 09/128,709  
 PRIOR FILING DATE: 1998-08-04  
 PRIOR APPLICATION NUMBER: US 60/054,645  
 PRIOR FILING DATE: 1997-08-04  
 PRIOR APPLICATION NUMBER: US 09/130,491  
 PRIOR FILING DATE: 1998-08-06  
 PRIOR APPLICATION NUMBER: US 60/054,966  
 PRIOR FILING DATE: 1997-08-06  
 PRIOR APPLICATION NUMBER: US 60/058,108  
 PRIOR FILING DATE: 1997-09-05  
 PRIOR APPLICATION NUMBER: US 09/388,280  
 PRIOR FILING DATE: 1999-09-01  
 PRIOR APPLICATION NUMBER: US 09/388,279  
 PRIOR FILING DATE: 1999-09-01  
 NUMBER OF SEQ ID NOS: 20  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 16  
 LENGTH: 551  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-802-582-16

Query Match 1.3%; Score 12; DB 12; Length 551;  
 Best Local Similarity 100.0%; Pred. No. 0.066;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 DIVTIPAGATNI 717  
 DB 324 DIVTIPAGATNI 335

Search completed: May 7, 2004, 12:06:24  
 Job time : 55 secs



```

Db      361 TVCDPSRSCSVIEDDGLQAAFTTAHLEIGHVFNMPHDDAKKACASLNGVTDGSHLMSLS 420
Qy      400 LNCGLPMSPCSAWVLTLLDGGHDCCLDAFGAALPLPTGLPGRMALYOLDQOCROIIFG 459
Db      421 LDHSPMSPCSAWVLTLLDGGHDCCLMDKQNFETKLPSPDLC--TLTDANRCQCTTFGE 478
Qy      460 DFRACPTNSADVCAQALWC--HTDGAELPLCHTKNGSLPMADGTPCGPHGLCSEGSCLPEER 518
Db      479 ESKICPD--AASDTTLTKCTSGGLVCQTKH--PFMADQTSCEGKMCVSGKCVNKT 534
Qy      519 VERKPVYDGMAMWPGWEGSRTCCGGVOPSHNECDPEPONGRCYCLGRRAKYOSCHT 578
Db      535 MKHFAATPHGSMGWPGWGDSCRTCCGGVQYTKRECONPVPKNGKTCBGRVRYSCNT 594
Qy      579 EECPP-DGKSPREDOCEKXNAYNTDMDGN--LLQWPKYAGVSPDRCKLFCRARGSE 635
Db      595 EDCPDNNGKTFREGECEAHNEFSKASF--GNEPTVEWTPKAGVSPKRCXCLTCEAKGIGY 653
Qy      636 FKVEPAKTIIDTLCGPETLACVNGCCYKAGCDHVVDSPKLDKCGYCGGKNSCRVSG 695
Db      654 FFVLQPKVVDGTPSPSPSTSVCGQCYKAGCDRIIDSKKFKDCGVCGANGSTCKXMSG 713
Qy      696 SLTPNFGYNDIVITIPAGATNIDVKORSHPGVQNDGNYLAKTADGGYLLNGMLAISALE 755
Db      714 IYSTRPGYHDIVITIPAGATNIEVKHNRGSRMNSFLAIRADGYILLNGMPTLSTLE 773
Qy      756 QDILVKGITLKYSGLATLERLQSFRLPEPLTVQLLTPGEVFPFKVKTFFPVNDVF 815
Db      774 QDLTKYGTALRYSGSSALERIRSFPLKEPLTIQVMV--GHALRPKIKFTYFM----- 826
Qy      816 SMGSKERATNIIQPLHAQWVLGDMECSSTCGAGMORTEVCRPDSGOASATCKAL 875
Db      827 ----KKKTESFNALPTF--SEWVIEHMGESCKTGGSGMGRVVOCCRDINHPSACEKEY 880
Qy      876 KPEDAKPCESQLCP 889
Db      881 KPASTRPCADLPCEP 894

RESULT 2
T00355
hypothetical protein KIAA0688 - human
C:Species: Homo sapiens (man)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00355
R:RishiKawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-837 <1SH>
A:Cross-references: EMBL:AB014586; NID:G3327189; PIDN:BA31663.1; PID:G3327190
A:Experimental source: Brain
C:Genetics:
A:Gene: KIAA0688
C:Superfamily: thrombospondin type 1 repeat homology
F:519-576/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match      38.0%; Score 1842.5; DB 2; Length 837;
Best Local Similarity 45.0%; Pred. No. 2.9e-117;
Matches 375; Conservative 130; Mismatches 244; Indels 85; Gaps 24;

```

```

Db      141 SLMWDGALLVLOYRGELHLQLEGGTPASAGPGA--HILKRKSPAGC---QGMWC 194
Qy      169 EYETGEQGRORGDHQEDSEESQEEAEAGSEPPPLGATS-----RTKRVSSEARFVET 224
Db      195 NVKA-----PLGSPSPRPRAKRFAASLSRFVET 222
Qy      225 LTVADAAAFYAGADLQNHILTLMSVAARIYKHPISIKNSILMAYVKULIYDEKMGPEVS 284
Db      223 LTVADDKKAAEFAGALIKTYLLTVAAAARAKRHSIRNPVSLVYTRVLITSGSEGPVG 282
Qy      285 DNGGLTLNFCWQRFRNPQSPRHEHYDTAILLRQNFCCQOEGCDTLGVADIGTICDP 344
Db      283 PSAQTLNFSFCAMQDGLNTPEDSDPDHDTAILLFRQDLCG--VSTCDTLGMADVGTVDP 341
Qy      345 NKSCSVIEDGLQAAHTLAHELGHVLSMPEDSKPCTRLPFGMG--KHVYMAPLVHLNQT 403
Db      342 ARSCAIVEDDGLQSAFTAAHELGHVFNMLHNSKPCISLNGPLSTRHYMAPVMAHVDPE 401
Qy      404 LMSPCSAWVLTLLDGGHDCCLDAPGALPLPTGLPGRMALYOLDQOCROIIFGDPFRH 463
Db      402 EPMSPCSAFITDFLDNGYGHCLDLPAPLHLPTFFGKO--YADNQCCULTEGDSRH 459
Qy      464 CENTSADVCAQALWC--HTDGAELPLCHTKNGSLPMADGTPCGPHGLCSEGSCLPEEYER 521
Db      460 CPQLPPP--CALMCSGHNG--HAMQTKHS--PMADGTPCGPAQACMGGRCLHMDQLD 514
Qy      522 EKPVDGMAPWPGWEGSRTCCGGVOPSHNECDPEPONGRCYCLGRRAKYOSCHTEEC 581
Db      515 FNIPQAGGMPWPGWGDSCRTCCGGVQPSRDCITPVVRNGKTCBGRVRYSCNTDSC 574
Qy      582 PP-DGKSPREDOCEKXNAYNTDMDGN--LLQWPKYAGVSPDRCKLFCRARGSEPK 637
Db      575 PMSALTFREGECAVYN--HRTDLKSPFGPMWVPRTGAVAPQDQCLTCARALGYTY 632
Qy      638 VPEAKTIIDTLCGPETLACVNGCCYKAGCDHVVDSPKLDKCGYCGGKNSCRVSGSL 697
Db      633 VLEPRVVDGTPSPSPSSVCGQCYKAGCDRIIDSKKFKDCGVCGANGSTCKXMSG 692
Qy      698 TPTNFGYNDIVITIPAGATNIDVKORSHPGVQNDGNYLAKTADGGYLLNGMLAISALIQD 757
Db      693 RKPRYGVNNAVITIPAGATNIIVRQGNPE--HRSYTLAKLPDGSYALNGEYTLMPSPD 750
Qy      758 ILVKGIT-LKYSGLATLERLQSFRLPEPLTVQLLTPGEVFPFKVKTFFPV 810
Db      751 VVLPGAVSLRYSGATASSETLSHGFLAQPLTVQL--VAGNDQDTRLRYSPFVP 803

RESULT 3
T21371
hypothetical protein F25H8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T21371
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19413
A:Accession: T21371
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: Clone F25H8
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19949
A:Accession: T24896
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <W12>
A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
C:Genetics:

```

A:Gene: CESP.F25H8.3

A:Map position: 4  
A:Introns: 31/1, 52/1, 135/2, 193/3, 216/1, 266/1, 495/2, 547/3, 564/3, 634/2, 744/1, 81

Query Match 30.4%; Score 1476; DB 2; Length 2165;

Best Local Similarity 34.1%; Pred. No. 7.4e-92;

Matches 327; Conservative 150; Mismatches 319; Indels 164; Gaps 29;

46 RLPGSAGELALHSAREKGFVRLAPDPSFL-----APEFIEHLGGS 88  
 96 RLQGVARD-----GHACHLRBSDAVYVHLHRMNOIPSHKXSPHSNSNFAM 148  
 89 GRATGGERGLRG-----CFPSGTVNGEPESLAIVSIC--RLSGSFLLDGEETIQ 137  
 149 VLYDSEEVKGGMSRDPDCIYRAHVGVGHQ--STVNLCDSEDLGYGMALPFGHITVE 207  
 138 P-----QAGAGSLAQPHRLQGWPGARPLRGGEWEVEIGEGORORGHQ--QDSSE 189  
 208 PISGNGTEHDGASRRHQLVRKFDPMHFKSF-----DHLSNSTVN 248  
 190 ESOEEAEGASEPPPLGATSRTRFV--SEARFVETLLVADASMAAFYADILQNHILTM 248  
 249 ETEETVATWQDQWEDVIERKARSBRANSMDHYEVLVADTKMVEYHGSLEDVLTIF 308  
 249 SVAARIYKPSIKNSINLAVYKVLVEDEKWEVSDNGELTLRNCMQRFRNPQSDH 308  
 309 STVASIRHOSLRASINVVVVKLIVKTENAGPRITQNAQOTLQDFCRMOQVYNDPDS 368  
 309 PEHVDITLITLRFQGOEGLCDTLGVADIGTICDPKSCSVIEDGLQAHITLHELH 368  
 369 VQHDHVALILTRKDICSGQCKDTLGLAELGTWCDMOKSCAILEDGLSAFTIAHELH 428  
 369 VLSMHPDSSKPCRLFGPMGX-----HHVAPLFLVHLNQLPWSPCGA 411  
 429 VFSIFPHDDEKCS--TYPVNVKVCFKQSTKEDTKTOFQNNFHIMAPLTLEYNHPWSWSPCSA 487  
 412 MYLTELIDGHHG--DCILDAFGAAL--PLPTGLPEBMALYQLDQCCROIFGDPDRHCP 465  
 488 GMLERFLENRRGQTCFLPDQVEHRYEDVFNDEPKK--YAHQCKVFVPSAELCP 545  
 466 NTSADVCAQIMCHT--DGAEPILCHTKNGSLPMADGTPCGPH--LSESGCT--DEEYV 519  
 546 ---TMRFCRLMCAITFGSGMGRTQH--MPMADGTCDBSRSMFGHAGCAVLALESYL 600  
 520 ERKPVVDGKAPFPMGEGSCRTGCGGVQSHRECKDPEPONGRYLGRARATGCHH 579  
 601 K-----IDGQWGDWRSWEGCSRTGCGGVQKLRDCSPKPRNGKTCVGGREYRSCNTQ 655  
 580 ECPEDGKSFREOQCEKTN--AYNYTDMGNILQVFPKXAGVSPDRCKLFCRARGRSEFK 637  
 656 ECPHDTPRYREVQSEFNKNDIGIGVASTNTMHVPKXANVAPNERCKLYCRLSGSAFY 715  
 638 VFEAKVIDTLCGETTALCYRGCCVAGCDHVVDSFRKLDKCGVCGKANSCKRYSGSL 657  
 716 LLRDKYVDGTIPCDNRGDDICVAGACMPAGCDHQLHSLTRDKCGVCGGDSCKVYKTF 775  
 698 TPT--NYGNDIVTIPAGATNIDVQRSHPGVONGNYALKTADGQYLINGNLAISAIQ 756  
 776 NEQGTFGYNEVWKIPASANIDIRQKGYNNKEDNYLSLRANGETLNGHFCVGLARQ 835  
 757 DILVGTILKYSGSIATLERLSFRPLPEPITVOLLTPGSEVFPKRYKTYF--VPNDY 813  
 836 QIAQDVTLEYSGSDAIIERINGTPIRSDIYVHLVSGSH--PPDISYEMTAAPVNAV 893  
 814 -----DSM--QSSKERATYIIQPL----- 832  
 894 TRPISALYLRVTDYTWECRACRGGQSQKLMCLDMSHTRQSHDRKQCVLVRKQATRY 953  
 833 ---LHAQWVLGWSSEGSSTGAGMQRRTYECRPSG--QASAT--CNKALPEKAKPC 883  
 954 CNIDCSTRWITEVSSCSAKCGSGQKQKQVACVMEGDRQTPASEHLCDENSKPSDIASC 1013

RESULT 4

T47158

hypotheoretical protein DKFZp762C110.1 - human (fragment)

C/Species: Homo sapiens (man)  
C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000C/Accession: T47158  
R./Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.submitted to the Protein Sequence Database, March 2000  
A/Reference number: Z24379A/Accession: T47158  
A/Status: preliminaryA/Molecule type: mRNA  
A/Residues: 1-550 <AAA>A/Cross-references: EMBL:AL162080  
A/Experimental source: adult melanoma (Mewo cell line); clone DKFZp762C110C/Genetics:  
A/Note: DKFZp762C110.1

Query Match 29.8%; Score 1444; DB 2; Length 550;

Best Local Similarity 50.6%; Pred. No. 2.1e-90;

Matches 254; Conservative 92; Mismatches 134; Indels 22; Gaps 10;

391 HVMAPLFVHLNQLPWSPCSAHYLTLLDGGHGDCLDAPGALPLPTGLPGRMALYQLD 450  
 11 HMMASMLNLSHQPSWSPCSAAMITSFLDNHGGCLMDKPNPIQLPDDLPG--TSYDAN 68  
 451 QCCROIIFGPPDFRHCPTNSADVCAQLWC--HTDGAELPLCHTKNGSLPMADGTPCGPHLCS 509  
 69 ROCQPTFEEDBKCPD--AASTGTLWCTGSGGLVYQTH--FPMADGTSCEGKWC 124  
 510 EGSCLPEEVEVERPKVVDGKAPFPMGEGSCRTGCGGVQSHRECKDPEPONGRYLCLGR 569  
 125 NKCYNKTRDKHPTLPFGSGMGMPGDCRITGCGGVQYIMRCDNPVFNKGKXCGK 184  
 570 FAKYOSCHTECP--DGKSFREOQCEKNAVNTDM--DGNILQVFPKXAGVSPDRCKL 627  
 185 RRYRYSCLIEDCPDNGKTFREOCEAHNEFSKSFSGSPAVEMI PKXAGVSPDRCKLI 244  
 628 CRARGRSEFKVFEAKVIDTLCGETTALCYRGCCVAKGCAHVVDSPKLDKCGVCGKG 687  
 245 COAKIGYFVLQPKVVDGTPCSPSTSVQGGCVAGACRIITDSKKKFKCGCGGNG 304  
 688 NSCRVAGSLPTPNYNDIVTIPAGATNIDVQRSHPGVONGNYALKTADGQYILNG 747  
 305 STCKKISGVSIAKGYHDITITPGATNIEVKQNRGSHNNGSFLAIKADGYILNG 364  
 748 NLAIATIODILVKTTILKSGSIATLERLSFRPLPEPITVOLLTPGSEVFPKRYKTYF 807  
 365 DYTSLTLEODIMYKGVLRYSGSSAALERIRSFLEKPLTIOVLTV--GNALRPKIKITY 423  
 808 FVPNDVDSMOSKERATTNIIQPLHAQWVLGWSSEGSSTGAGMQRRTYECRPSGQA 867  
 424 FY-----KKKKEFPNAI--PTFSA--WVLEWGEKSKCELGWQRLVLCRDINGOP 471  
 866 SATCNKALPEKAPKESQLCP 889  
 472 ASECAGEVPASTRPCADHPCP 493

RESULT 5

T18517

procollagen N-endopeptidase (EC 3.4.24.14) I - bovine

N/Alternate names: procollagen N-proteinase  
C/Species: Bos primigenius taurus (cattle)C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C/Accession: T18517R./Collige, A.; Nuegens, B.V.; Lapierre, C.M.  
submitted to the EMBL Data Library, February 1996A/Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.  
A/Accession: T18517A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA

A:Residues: 1-1205 <COL>  
 A:Cross-references: EMBL:X56389; NID:e990769; PID:e228215; PIDN:CAA65253.1  
 A:Experimental source: skin  
 C:Genetics:  
 A:Gene: PC I-NP  
 C:Function:  
 A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to  
 C:Keywords: hydrolase; metalloproteinase

Query Match 22.7%; Score 1102; DB 2; Length 1205;  
 Best Local Similarity 30.9%; Pred. No. 9.6e-67;  
 Matches 316; Conservative 137; Mismatches 379; Indels 190; Gaps 41;

```

QY 3 PAPAAPMWL-PELLLILLPL- -LARGA- -PAPAGGQASELVPT- - - - - 46
DB 4 PAGAAGRLCPALILLPLPADARLAAADPPGGQGGARILLAVPRDAQRLV 63
QY 47 - - - - -LPG- - - - -SAGELALHLSAFGKGFVRLAPDPSFL 76
DB 64 SHVSAAPAPAGVTRRAPAQIPGLSGSSEDDPGRIFFVTVFGDHLRLRNRLV 123
QY 77 APEFIERLGGSGRATGGERGIRGCFPGTNGEPEBIA-AVSLCRGLSGFLIDGEEFT 135
DB 124 ARGATVEMQESG-ATRVEPILGCTLYGDVAGLASSVALSNODIAGLIRMEEEFF 182
QY 136 IOP--QAGGSLAQPHLQ--RMCPAGARPLPRGEWEVETGEGORQERGDHDSSEES 191
DB 183 IEPLEKGLAAKAEQGRVHVYHRRPTSRPPLGGPQALDGG- - - - -ISADSLDST 233
QY 192 QE- - - - -BEAGASRPPPLGATSRKTFVSEARP-VETLLVADASMAFYAGD-LONHI 244
DB 234 SRALGVLEBRVNSR- - - - -RMRHAAADDYINIVLLGVDSDVQFHTGHVQKYL 285
QY 245 LTLMSVARIKXKPSIKSINLMVVKLIVEDEKMGPEVS- - - - -DNGSLTIRNF CMMQRR 300
DB 286 LTKMIVNEIYHDESLGHINIVLVIRILL- - - - -SYGKMSLIEICNPQSLNCRWAYL 342
QY 301 PNOQDRRPHRYDTAILLTRONFCQGEGLCDLTVADIGTICDPKSGSVIEDGLQAH 360
DB 343 QOKPDTDEYHDHAILFLTRDF--GPSGM--QGAIVTGMCHPFRSCTLHNEGFFSAF 398
QY 361 TLAEHGHVLSMHPDSDKPTRLFGPMKSHVMAFLFVHLNQLTPMSPCSAMYLTLLDG 420
DB 399 VVAHETGHVLMGHEHGGQ--NRGDEVRILSGIMARLVGAARHFRWMSCSQGEISRYL- 454
QY 421 GHG-DCLLDAP- - - - -GAALPLFTGLPRKMAIYQLDQCRQIFGDFRRCNTSAQDVCAQ 475
DB 455 -HSYDCLRDDPFTDMFALPOLPLG- - - - -YSNMEOCRFDGAGYMCATFRFTDPCKQ 508
QY 476 LMCHTDGAEPICHTKXGSLPMADGTPCGGHLCEGSGC--LPEEVEPRKPVVGGMAFW 533
DB 509 LMCSPHDNPFCKTKKG--PLIDGTMCAPGKCHFCFGHCLWLPDLIK- - - - -DGNMAM 561
QY 534 GFWGCSRTGGGVOPFSHRECKDPEPONGARYCLGRARAYOSCHTEECPPDGKSFREOC 593
DB 562 SEFGCSRTGTVFRTRQCDNPHANGARTCSGLAYDFOLCNSQDPCDPLADFREBOC 621
QY 594 EKXNAY-NYTDMDGNLQWPKYAGVSPRDRCKLFCRARGSEPFVFAKYLIDGLCG-P 651
DB 622 RQMDLYFEHDAQHH--MLP--HEHRDAKERCHALCESKEKGEVVMGRMHHDTGRCYK 677
QY 652 ETLAICVRQCVKAGCDHYVDSFRKLDKGVCGGKNSCRKYSG--SLTPNYGNDIVT 709
DB 678 DAFSLCVRGDCRKGVDGIVIGSSKQEDKGVCGGNSHCKYKGFSSPKKLGIYKFE 737
QY 710 IAGATNIDVKORSHGVONDGNIALAKAD--GQVLLANGNLASIAEDDILYKGLIKYS 768
DB 738 IPAGASHLLIQE- - - - -ADTSHHLAVKMLETKILNEENDVDPSKTFIAMGVMEYR 792
QY 769 GSIALTERLOSPFLPEPLVOLTVPGEVPPKVKYTFVYND- - - - -VD- - - - - 814
DB 793 DEGR-ETLQTMGFLGITTV--LVIPEDARISLYKMYIHEDSLANDNNVLDDSVG 849

```

```

QY 815 - - - - -FSMQSSKERATTNIIQPLHLAQ 836
DB 850 YEMALKWSPCSPGCGSGQFTKYGCRRLRDLHKVHVRGCDVSXPKAIRTCNPOESQ 909
QY 837 - - - - -WYLGDMSECSSTGC-AGWQRRTYECDDPSQASAT- - - - -CNKALKPEPAKCEQL 887
DB 910 PAVVWGEWPPCSCSGRTGMQVSVRQVQPLHNNTTRSVHTKHKND- - - - -RPGRRACNREL 968
QY 888 CP 889
DB 969 CP 970

```

## RESULT 6

C89114  
 Protein C37C3.6a [Imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: C89114  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio  
 A:Reference number: A75000; NUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.banger.ac.uk/Projects/C\_e  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an  
 A:Accession: C89114  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1558 <STO>  
 A:Cross-references: GB:chr\_V; PIDN:AAC25867.1; PID:G3294501; GSPDB:GNO00023; CESP:C37C3.  
 A:Gene: C37C3.6a  
 A:Map position: 5

Query Match 11.6%; Score 562; DB 2; Length 1558;  
 Best Local Similarity 34.9%; Pred. No. 6.7e-30;  
 Matches 129; Conservative 51; Mismatches 136; Indels 54; Gaps 14;

```

QY 528 GGMWPMGPGSGRTGGGVOPFSHRECKDPEPONGARYCLGRARAYOSCHTEECPPDGKS 587
DB 77 GMMGPMVPEBEGSRSGGGVQLKQCSGD- - - - -CTGASVYISGNLMAC-ESGTD 127
QY 588 FREOQEKXNAYNYTMDGNLLQWPKYAGVSPRDRCKLFCRARGSEPFVFAKYIDGT 647
DB 128 FRAEQCKEND--EALDGNVHKMP-YKG--KNKCELVCPBESGNFYKADRVADGT 180
QY 648 LCGPFTLACVAGQCVKAGCDHYVDSFRKLDKGVCGGKNSCRKYSGLTPTY--GYN 705
DB 181 KCDKSKNDICVDGECUPVQCDGKLSLFPDKGCKDGDGSLTKTIEGRFDRNLSPGH 240
QY 706 DIVTPAGATNIDVKORSHGVONDGNIALAKADGQVLLANGNLASIAEDDILYKGLT 765
DB 241 DIILPEGATNIIQE- - - - -ARKSTNNLALXNGSDHFLYNGN-GLIQYEKEVEVGITF 294
QY 766 KYSGLATLERLOSFPFLPEPLVOLTVPGEVPPKVKYTFVY- - - - -NYVDSMGSSKER 823
DB 295 VYDD--AEPETSAQSPLEBELVALLFRKGS-RDTALTYEISIPLEBEVDY- - - - - 343
QY 824 ATNIIQPLHLAQWVLDGDMSECSSTCGAGMORRYECRDPGQASA--TC--NKALKPE 878
DB 344 - - - - -MYKEDMTFPCSVSGKGVQTRNLVCLDGKKGARVEDLCEENATKE 391
QY 879 DAKPESQLC 888
DB 392 FEKSCETVDC 401

```

## RESULT 7

T34395  
 Hypothetical protein C37C3.6b - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 01-Dec-2000  
 C:Accession: T34395; T34394

R:Geisel, C.; Bradshaw, H.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: The sequence of C. elegans cosmid C37C3.  
 A:Reference number: 221518  
 A:Accession: T34395  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2167 <GEI>  
 A:Cross-References: EMBL:U64857; PIDN:AAC25868.1; GSPDB:GN00023; CESP:C37C3.6b  
 A:Experimental source: Ectrain Bristol N2; clone C37C3  
 A:Accession: T34394  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1555 'SKF' <GS2>  
 A:Cross-References: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a  
 A:Experimental source: Ectrain Bristol N2; clone C37C3  
 A:Genetics:  
 A:Gene: CESP:C37C3.6b; CESP:C37C3.6a  
 A:Map position: 5  
 A:Introns: 32/3; 104/2; 156/2; 207/1; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 1556/

Query Match 11.6%; Score 562; DB 2; Length 2167;  
 Best Local Similarity 34.9%; Pred. No. 1e-29;  
 Matches 129; Conservative 51; Mismatches 136; Indels 54; Gaps 14;

QY 528 GGNAPGMPGECSTCGGVOFHSRECKDEPQNGRGYCLGRRAKYOSCHTECPDGGKS 587  
 DB 77 GNMGPMPBENCSSCGGVOLEKQCSG-----CTGASVRIYISCLNAC-BSGID 127  
 QY 588 FREQCEKRYAANYTMDGNLQWPKYAGVSPDRCKLFCRARGSEKVFPAKYIDGT 647  
 DB 128 FRAQCEKQND---EALDGNHMKWP-YKG---KMKCELCKPESGIFYKMAKDVDTG 180  
 QY 648 LCGEETLACVCGQVACGCHVDVSPRLDKCGVCGKSGNSCRKYSGLTFPNY---GVN 705  
 DB 181 KCDKSKNDICVGGCLPVGCDGKGLSSLPKDCCKCGDGSCTCTTGRDERNLSPGTH 240  
 QY 706 DIVIPAGATNIDVKORSHPCQNDGNVYALKTADGQYLINGNLASAIQDILVKGTL 765  
 DB 241 DIILPAGATNIIQIE---ARKSTNLAKNGSDHFLNGN-GLIQVEKEVEVGITP 294  
 QY 766 KYSSSIATLERLQFRPLPEPLTVQLLTVGGEVPPPKYKTFVPV---NDVDFSMQSKER 823  
 DB 295 VYDD--KEPFLSAGQPSSEELTVALLFRKGS-RDTAKYEFSTPLEEVDY----- 343  
 QY 824 ATTNIQPLHAQVLAGMSECSSTCGAGQRTVECRDPSGQASA---TC--NKALKPE 878  
 DB 344 -----MYKFDNWTPCSVSCGKGYQTRLYCIDGKNKGRVDDLCENNAITKPE 391  
 QY 879 DAKPCEGOLC 888  
 DB 392 FEKCEETVDC 401

RESULT 8  
 T18856  
 angiotensin inhibitor homolog - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T18856; T24653  
 R:McMurray, A.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: Z19031  
 A:Accession: T18856  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1444 <W1L>  
 A:Cross-References: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1  
 A:Experimental source: clone C02B4  
 R:McMurray, A.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: Z19917  
 A:Accession: T24653

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1444 <W12>  
 A:Cross-References: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1  
 A:Experimental source: clone T07C5  
 C:Genetics:  
 A:Gene: CESP:C02B4.1  
 A:Map position: X  
 A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 56

Query Match 10.9%; Score 529; DB 2; Length 1444;  
 Best Local Similarity 24.1%; Pred. No. 1.e-27;  
 Matches 236; Conservative 119; Mismatches 366; Indels 258; Gaps 48;

QY 41 LVVPTRLGSGELALHLSAFGKGVFLAADDSTLAE-FKIEHLGSGRAT-GESEGL 98  
 DB 31 LIEATRHPLKGNLTKMFTAMNDTYHLNLRKSRIVSPHITSVHGDDVTTAGLNDY 90  
 QY 99 RCGPFGTVNGEPESIAAVSLCRGLSGSFLLDGEFTIQ--POGAGSLAOPHRLQRMGP 156  
 DB 91 ECGHYQGEVYKSGNKKALISDCALMSIVMEDHFLVQLPK-----RVNHLQX--- 140  
 QY 157 AGARPLPRGPEWETVGGQROEGRDHOE---DSEESQEBEABGASEPPPLGATGRT 212  
 DB 141 -----EHLVYKRSAGLLTNAESKIREIITRLOEBOESFCDTSQDDPMTTIPAHNH 194  
 QY 213 KRFVS-----BARFY-----ETLVADASNAF-----YGADLQNH1-----TLMSV 250  
 DB 195 NTITPTSQALDSSTFPNMDDPTLEIGLFLDSKLFENHERYIQDAEQHLEFSLALINN 254  
 QY 251 AARIYKHSIKNSINLWVYKLYVEDEKMBE-----VSDNGGL--TLRNFQWQRF 301  
 DB 255 VAVLYQDPTLFPNLDIVYRY-----EWMRTQPSALSTGVHKNCAQGLLAFCRYQAHM 309  
 QY 302 NQPSD-RHREHYDTALITRQNFQGBGLCDTLGADIGTICDPKSGSVIEDEGLQAH 360  
 DB 310 NFGTLDLDMNHHDHVLILGYDI-YHTTTSVAGAPARAKCDBLPACSLVEGLHLSGF 367  
 QY 361 TLAEHLGHVLSMFHD-----SKPCTRLFPGMKHNVAPLFLVNLQTLPMSPGSA----- 411  
 DB 368 VLAHEMGNHMGVHDGVQNCNKGCCLSMAVNAQAKT-----TMSDCSVREFN 415  
 QY 412 MTLTELDGGHDCILDA-EGAAFLPTG-----LPRMALYQDDCCROFEGDFR-H 463  
 DB 416 AFLQLDSGRNCLRDASFGI--ISTNHLSDLRLPQR--FLADQCCSTFWGADYVE 470  
 QY 464 CPTSA-QPVCAQLMCHTDGAFLCHTRKNGSLPMADGTPCGGHLCSEGSCLPEEVEVP 522  
 DB 471 INGKAMDICRILMCGNSG-----TSTAHPLLEGSGCANKKCHGQCTHMTFFGLTP 525  
 QY 523 KPVVDGMAFNG-----PMGECSRTCGGVOFHSRECKDEPQNGRGYCLGRRAKYOSG 576  
 DB 526 VP-IDGESEWGAKEGCEPIQCAVSGSITYQGGHRCDVNPAFNNGGTTC----- 574  
 QY 577 HTECEPPDGKSPREOC--EKTNAANYTMD-GNLQWPKYAGVSPDR----- 623  
 DB 575 -----EGANIRGLVCGATISNCLGTFREBRGNKISIKXDPRKPOQLGSEFHSHT 627  
 QY 624 --CKLFCRARGSEKVFPAKYIDGTLGPELTALCVRGQCYKACDHV-----DSEF 675  
 DB 628 QCRVYCHLIG-SELIRKNGQFPDGTGCFD--ACVGGQCCIALSGCKALVEQDEDPFR 684  
 QY 676 -----KLDKGV-CGGKNSCRKYSGLTFPTNGYNDIVITPAGATNIDVQ 721  
 DB 685 IGRSVHQMEESSMSSECSVSCGLGREVREKCS-----SRKQGVSE 729  
 QY 722 RSHP--GVQNDGNVYALKTADGQYLINGNLASAIQDILVKGTLKTSGSTA---TLER 776  
 DB 730 ESRPCEGVLRCDEEF-----GEW-----KEMGSCSEKCAAGV 761  
 QY 777 LOSFRPLPEPLTVQLLTVGGEVFP-----KVKTTFVP 810  
 DB 762 QKRFRC--LTDQCSKHLQBERPCDNEGCWTNMDWSSCGSGGRRRYIRKCLDDK 818

QY 811 NDVDFWSSQSKERATTTIIGPLHAQWVLDGMSCCSTCGAGMORRTVECDRDSGQASAT 870  
 DB 819 CDGG---DLEESCNT---QXCISQSW--GDWLPSCVSGIIGFOIRERLC---DGLICAT 867  
 QY 871 CNKALKEPAKPCESQOLCP 889  
 DB 868 ANK-----QARTCNQOCCP 881

## RESULT 9

116892  
 Hypothetical protein T19D2.1 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C/Accession: T16892

R/Bentley, D.  
 Submitted to the EMBL Data Library, December 1995  
 A/Description: The sequence of C. elegans cosmid T19D2.  
 A/Reference number: Z18599  
 A/Accession: T16892  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-860 <BEN>  
 A/Cross-references: EMBL:U42846; NID:G1125809; PID:G1125810; PIDN:AAA83600.1; CESP:T19D2  
 C/Genetics:  
 A/Gene: CESP:T19D2.1  
 A/Introns: 56/3; 96/1; 171/3; 199/3; 245/2; 293/2; 335/3; 395/3; 426/3; 484/2; 505/3; 58

Query Match 9.8%; Score 473.5; DB 2; Length 860;  
 Best Local Similarity 23.9%; Pred. No. 3.3e-24;  
 Matches 190; Conservative 109; Mismatches 284; Indels 213; Gaps 44;

DB 182 DH--QEDSEESQEEAAGASEPPPLGATSRTRKRVSEAR-----FVETLLVAD 229  
 DB 90 DHEOCFDANDPYPDRR-----ELAMKSMILRMVMDIRAPRRDLITVELAVFAD 142  
 QY 230 ASM---AATYGADLQNHILT-LMSVAARI---YKHSIKNSINLMVVKYLIVD--EKV 279  
 DB 143 DAMMDHFKKMYGKALENNMTFIMAVVNNIDVLTORLLORINIKIVREILKNIPHLM 202  
 QY 280 GREVDNGGL--TLNFCNMQRFPNOPSRRHEHYDTAIIITRONFGQEGELCDTLGVAD 337  
 DB 203 NARKISNDVRLDAFCQYQNEINPRVADPRHMDALLFSGYDL-HRNGVKTVAGYAP 261  
 QY 338 IGTICDPNKSQSVIIDEGLQAAHTLAHELGHVLSNPHDSDKPCR--LRGP--MGKH 391  
 DB 262 VKGMSGVRSCTINIGLDFGSVFVVTHEMGSHTGMDNECDLRCCIMSPVSGSKTH 321  
 QY 392 VMAPLFVHNLQTLFWSPCSAWYLTLLDGGH-GD-----CLLDAPGALPL----PTGL 440  
 DB 322 -----WSQCSVNMATFTV--GHIGDFRPPNLOQASNAEQGMVAFKSESP 365  
 QY 441 PGRMALYQLDQCRQIFGDFRH--CENTSADVCQAQJWC-HTDGAEPDCHTKNGSLPWA 497  
 DB 366 PQQ--LFTLDEQCELFHGCWGHLEKDGQIMQNICQWVWCNGEGVIRTAH-----PAL 417  
 QY 498 DGTGPGPGHLCSESGCLPEEYERPKPVVDGMAWGMWGSCSRICGG-----G 546  
 DB 418 EGTGGGFMITCRQGCQVSSQLMR---VTVGSMSTWN--DRPAPTCGRCGQCEIRQIR 472  
 QY 547 VQFSRECKDPEFONGRXYCLGRARYQSCHTCECPDQ--KSFREQCEKY---NAVNYT 602  
 DB 473 IMRSIRQCNPSNNGARPCQGDARQVCHRDVNGDSIERVATRVCSRLDEALAPNT 532  
 QY 603 DMDGNLLQWVPKYAGVSPRDRCKLFCRARGRSEFFKFAKVIDGLCGPELTALCVBQC 662  
 DB 533 IISGEMQF-----EQAMCKIWCLISGSTNITIV--SNFPDGAFCGG--QYIKGEC 581  
 QY 663 VVAGCDHVDPKDKGVCGGKNGSKVAGSLTPNYGND-----IYTIIP-- 711  
 DB 582 -----RPL-----LQGS-----TLTAISEADCLPLSLVQTTPPM 611

QY 712 -----AGATNIDVKQSHPGVQNDGNLALKTADGGQYLNGNLAIABODILVKG 762  
 DB 612 PHVSVDPFAGKTNPYKHKKTPELNEWSGV-----SWASECVTVDCHTQG 657  
 QY 763 TILKSSGSIATL-----ERQSPRL--PEPLTVQLTVGGEVPPKVVYTFEPNDV 813  
 DB 658 VKVRVRCLAGVAGALREPRQPTRPCTGSERPILT---SPQOTF--RRFIAPLRV-- 710  
 QY 814 DFSWQSKERATTTIIGPLHAQWVLDGMSCCSTCGAGWQ-RTVECDRDSGQASATCN 872  
 DB 711 -----RQTNMIRKVDH--W--GPWSACSVTCGTGTLRRRENC---IGQBCA--- 751  
 QY 873 KALKPEDAKPCESQOLCP 888  
 DB 752 -----ETGFCVMQSC 761

## RESULT 10

T00260  
 Hypothetical protein KIAA0605 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 21-Jul-2000  
 C/Accession: T00260  
 R/Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.  
 DNA Res. 5, 31-39, 1996  
 A/Title: Prediction of the coding sequences of unidentified human genes. IX. The complete  
 A/Reference number: Z14086; M01D:98290545; PMID:9628581  
 A/Accession: T00260  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-951 <NAG>  
 A/Cross-references: EMBL:AB011177; NID:G3043733; PIDN:BA25531.1; PID:G3043734  
 A/Experimental source: brain  
 C/Genetics:  
 A/Note: KIAA0605  
 C/Superfamily: thrombospondin type 1 repeat homology  
 F/46-106/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 9.0%; Score 438; DB 2; Length 951;  
 Best Local Similarity 31.7%; Pred. No. 9.8e-22;  
 Matches 113; Conservative 45; Mismatches 124; Indels 74; Gaps 14;

DB 526 VDGG-----WAPWPGECSSRTGCGGVQSFHRC-----KDPFONGRXYLGRARYQ 574  
 DB 39 LRGGTDAIFWNGEWTKTAFSRSCGGVTSQERHCLQQRKSVGPGRCTGTGSKRYQ 98  
 QY 575 SCITECPDQKSFREQCEKYNAYTMDGNLLQWVPKYAGVSPRD-----RCKLF 627  
 DB 99 LCRVQCECPDQKSFREQCVSNHIVY--NRTTHQMKPLV---PDDVHHSKPCDLH 151  
 QY 628 C-RARGSEPKVFEAKVIDGLCGEPLT-AICVRQCVYAGCDHYVDSFRKLDKGVCGG 685  
 DB 152 CTTVDQQRQLMW---PARQTSCKLTLRGVCVSGKEPFGCDGVLFTHTLDKQGIQCG 208  
 QY 686 KNSCKRNVGSLTPN--YGYNDIYTIIPAGANIDVQKSHGVQNDGYALAKTADQY 743  
 DB 209 DGSSCTHVGNYRKGAAHGYSLVTHIFAGARDIQVER-----KKSADVLLADEAGY 263  
 QY 744 LINGNLAIABODILVKTILKYSGSI-----ATLERLOSFPRLPEPLTVQ----- 790  
 DB 264 FRNGNKKVDS-PNNFNIAGIVVYKRRPMDVEYGEIYVIAQGFPTQGLVWVWNGKSP 322  
 QY 791 -----LTVPEGVFPKVKYTF-----FVNDVDFWSSSKER 823  
 DB 323 SITFEYTLLOPHESRPOPYYGFSASQGLDAGLGMFLPHNGSLYQASER 378

## RESULT 11

T11576  
 Hypothetical protein F08C6.1 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C/Accession: T11576





```

QY 565 YCLGRAPKQSCHECEPDEGKSFREQCEKXNAVNTYMDGNLLQWVPKYAGVSPDRD 624
DB 472 LCRGR-----TSECD-----VPEX-----485
QY 625 KLFCAARSEPFKVEAVIDGTLGPEPLAIYVRGQC--VKAGDHYVDSPRKL-----677
DB 486 -----CNSSQPCQDPVFTQNGYPC-QNNKAYCTNGKQYDQCVYFSGKAPAPDC 539
QY 678 -----DKGVCCKGKNSCRK-----VSGSLPTFN-----YG-YNDIVTIPAGAT- 715
DB 540 FLEVNSKDRFNGCNCFSGNEXKCATGNALCKLQCENVGELPYFGVLPALIQPSSGTX 599
QY 716 --NIVKORS--HPGVNDG 731
DB 600 CWGVDFOLGSDVDFDGMVNEG 620

```

## RESULT 13

monocyte surface antigen MS2 precursor - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 26-Aug-1999  
 C/Accession: A60385  
 R.Yoshida, S.; Setoguchi, M.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
 Int. Immunol. 2, 585-591, 1990  
 A/Title: Molecular cloning of cDNA encoding MS2 antigen, a novel cell surface antigen  
 A/Reference number: A60385; PMID:91197896; PMID:1982220  
 A/Accession: A60385  
 A/Molecule type: mRNA  
 A/Residues: 1-826 <YOS>  
 A/References: EMBL:X13335  
 C/Superfamily: mouse melanin alpha; disintegrin homology  
 C/Keywords: glycoprotein; surface antigen; transmembrane protein  
 F/1-14/Domain: signal sequence #status predicted <SIG>  
 F/402-484/Domain: disintegrin homology <DIS>  
 F/659-683/Domain: transmembrane #status predicted <TM>  
 F/330/Active site: Glu #status predicted

Query Match 6.4%; Score 309.5; DB 2; Length 826;  
 Best Local Similarity 22.4%; Pred. No. 4,5e-13;  
 Matches 171; Conservative 85; Mismatches 292; Indels 215; Gaps 36;

```

QY 14 LLLLLLLPLARAPAPPAQASLVPTPLPGSAGELAL-----HISA 60
DB 4 LMLSLVLTWPAY--APGPPLPHVKQYEVWPRRLASRSRALPSHWQYFESLYALGT 61
QY 61 FQKGVFLIAPDPSFLAPFKIERLGGSG-RATGGERLRCGFPSTNGEPESLAVSL 119
DB 62 SHVFTLHKKRDLIGSSYETTYTSANSGSEYTELQDHLQGHVEGREGSASIST 121
QY 120 CGLSGSFLIDEEFTIOP-----QAGGSLAOPHLRQWGPAGARPL-----PRGEW 168
DB 122 CAGLGGFFRVGSTVHLIEPLDADEGQAMVQAKHLQKAGTCVQDNTNLNDLGPRALE- 180
QY 169 EYETSGQRQERGDQDESEESQEBEAGASEPPPLGATSRTRFVSEARFETLLVA 228
DB 181 -----YRAOPRMKLIREFRYELVVA 204
QY 229 DASMAFYGA--DLQNHILTMSVARIYKPSIRNSINLWVVKLYIYEDKMGPE--V 283
DB 205 DQGFQKQKSRRAVQRYLEVNVHVDKLYQELSF-----RVLVGLGLIMWKDKYI 255
QY 284 SNGGLITLNFQCMQRRFQPSDRPHEDTAIILTRQFCQGESLCTTLGVAADIGTICD 343
DB 256 SRYAVNTLEFLSMBEQNLQ--GQHP--HDNVQILTGVDFIG--STVGLAKVASLCS 306
QY 344 PKKSGCVID---EGLOAHLTAHELGLVSPHDSKP-CTRLPGMKSHVVAFLPVH 399
DB 307 -RHSQAVNDHKSNIQVASTVAHLGLNLSHEDDIPGC--YCFPRPGGGGIMTES 362
QY 400 INQTLIP--WSPCSAMYLTELLDGGHDCILDAFGAALPLPTGLPERMALYQLDDQ-----451

```

```

DB 363 IGSKEPRIRSCSKIDLESFVTKPQTGCLTNVPVNRFLVGVGCVNLFVHEGQDCGTP 422
QY 452 -----OQRQIFGPD-----FRHCPTSAQDVC-----AQLWCHTGAEL 486
DB 423 QDCQNPCCNATTQQLYKGAEBGASGTCHECKVKPAGVETRLSKDCDLEEFQ--DGRKPT 480
QY 487 CHTKNSLP-----WADTPCGRGHLCBSGSCPESEVEPRKRVVNGMAPMGMBECST 542
DB 481 C-----PEDAFQQGTFC-PGYCPCDSCTTLAQCCDL-----WGP-----516
QY 543 CGGQGVQSHRECKDPEPQNGRGYCLGRARAKYQSCHECECP--DKSF--REQCEKXNA 598
DB 517 -----GARVADSCVTFSPQCGNRMVSGRLNRCAL--549
QY 599 YNTYMDGNLLQWVPKYAGVSPDR--CLT-----FCRARGR--SEFKVFEAVIQTICG 650
DB 550 --YCE-----GGQKPLERSFTFSNHHGVCHALGTGNSITDFEL-VLQGTIC- 593
QY 651 PETLACVAGQCVKAGCDHVVDSPRLDKCGVCGKGN-SCRK 692
DB 594 -EKGKVCMDGSCQDLRYVRSNCSAKNNHGVCHKRECHCK 635

```

## RESULT 14

hypothetical protein F53B6.2 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C/Accession: T22545  
 R/White, S.  
 submitted to the EMBL Data Library, October 1996  
 A/Reference number: Z19578  
 A/Accession: T22545  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 11059 <WLL>  
 A/References: EMBL:281086; PIDD:CA03121.1; GSPDB:GN00019; CESP:F53B6.2  
 A/Experimental source: clone F53B6  
 A/Genetic: C  
 A/Map position: 1  
 A/Intons: 38/3; 92/3; 131/3; 169/1; 236/2; 282/3; 349/2; 376/2; 420/1; 576/3; 759/3; 8

Query Match 6.4%; Score 309.5; DB 2; Length 1059;  
 Best Local Similarity 24.1%; Pred. No. 6.1e-13;  
 Matches 91; Conservative 45; Mismatches 113; Indels 129; Gaps 14;

```

QY 530 MAPGPMGSCSRCTCGGQVQFSHRECKDPEPQNGRGYCLGRARAKYQSCHECECPDQKSPR 589
DB 26 WAKNSPMSGCTKTCGGGVSRQLRLCLTSK-----CSGESVRFKVCQAKTCEKSRRLAR 78
QY 590 EQCEKXNAVNTYMDGNLLQWVPKYAG--VSPDRCKLPCPAR--GRSEFKVEAKYVD 645
DB 79 DTIC-----GGEIIVSRGCEVVCSSRLGAN-----FLWRVVD 112
QY 646 GTLCGPET-LAICVRQCVKAGCDHVVDSPRLDKCGVCGKNSCR-----KYSGSLT 698
DB 113 GTPCQATSRVAVSKSCQIVGCDGLISSFFPDAGVCGSGDPTDNDKFTMKYSEBT 172
QY 699 PTNYGVNDIVTIPAGATNIDVKKRSHPGVQNDGNTALATADQYLLNGNLAIASLEODI 758
DB 173 -----ACASND-----179
QY 759 LVKGTILTKSG--SIATLERLOSFRPLREPLTVQTLTVGSEVFPKVTYTFVVDVDF 815
DB 180 ---DIVDWSGASRSTASTS-----QPIVVCNATITGRVPEKL-----CADK 218
QY 816 SMOSSKERATTNIIQPLIHAQWTLGDMSECSSTCGAGMQRRTYECRDPGQAS-----AT 870
DB 219 LRKVEARPCPMILIC--SRWMAADWTECVPHCGGTGRKREYVCVOTAHNTVHVDPTE 275
QY 871 CNKALPEDAKPECSQLC 888

```

Fri May 7 12:18:02 2004

us-09-989-687-4\_1.rpr

Page 9

Db 276 CENGTBPAAENCSTSC 293

RESULT 15

G02390  
disintegrin-like metalloproteinase MDC15 (BC 3.4.24.-) - human  
C/Species: Homo sapiens (man)  
C/Date: 21-Dec-1996 #sequence\_rev: 06-Jun-1997 #ext\_change 31-Dec-2000  
C/Accession: G02390; PC4263  
R/Herrn, B.; Raines, E.W.; Ross, R.  
submitted to the EMBL Data Library, January 1996  
A/Reference number: H01157  
A/Accession: G02390  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-814 <HER>  
A/Cross-references: EMBL:U66005; NID:g1335871; PIDN:AA05112.1; PID:g1335872  
R/MKle, N.; Edwards, T.; Dallas, D.O.; Houghton, A.; Stringer, B.; Graham, R.; Russell, J.  
Biochem. Biophys. Res. Commun. 230, 335-339, 1997  
A/Title: Expression of members of a novel membrane linked metalloproteinase family (ADAM  
A/Reference number: PC4263; MID:97168971; PMID:9016778  
A/Accession: PC4263  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-461 <MCK>  
A/Experimental source: articular chondrocyte  
C/Comment: This protein is a membrane bound protein and involved in cell/cell and cell/m  
C/Superfamily: mouse metrin alpha; disintegrin homology  
C/Keywords: hydrolase; metalloproteinase; zinc  
F/420-503/Domain: disintegrin homology <Dis>  
F/348,352,358/Binding site: zinc (His) #status predicted  
F/349/Active site: Glu #status predicted

Query Match 6.3%; Score 306; DB 2; Length 814;

Best Local Similarity 25.1%; Pred. No. 7.6e-13;  
Matches 181; Conservative 75; Mismatches 273; Indels 192; Gaps 42;

QY 42 VVPTRLPGSAGELALHSAFGKGVTLRLPDDSLAPEKIERLGGSPAT-----GGE 95  
Db 60 VLQTSLP--EPRLKIKELDGDGSHLELQNR-LVP-----GRPTLWYQPDGT 105  
QY 96 R-----GLRGCFSGTGVNGEPESLAASLCRLGSGFLDGEFTIOPGAGGSLAQP- 148  
Db 106 RVVSEGHITLNCYQGRVRYAGSWISICCSGLRGLVVLTPRSTLIDQEGDLOGPPI 165  
QY 149 -HRLQWGPAGAPLPRGPWEVETGEGQRQERGDHDESEESQEEBAEGASEPPPLG 207  
Db 166 ISRIQDLHLPG-----HTCALSWRESVHTQTPPEHPLG 198  
QY 208 A--TSRTKRFVSARFVETILVADASMAAFYGLDNLHILTLMSVARI--YKHPKINS 263  
Db 199 QRIRRRRDVVTETKVELIVADHSEAKY-RDQHLNRTLEVALLLDTFFR----- 252  
QY 264 INLMVAVLIVDEKNGP-----EVSNGSLTLNFCNWRFRQPSDRPHHYDAIILT 319  
Db 253 ---LNVAVLVLGAEAWTQRLVEISPPAVTLNFTLWRRALLP--RLP--HDSALVLT 305  
QY 320 RQNPCCGEGLCDTLGVADITICDPNKSQVIEDEG---LQAHTLAHELGHVLSMHPD- 375  
Db 306 GTSFSG-----PTVGMALIONSIPDSGVMNDHSTSLGVASSTAHELGHSLGIDHDL 360  
QY 376 --DSKPTCLFLGPMGKHVVAFLVHLNQTLPWSPCSAMYLTEILDGHDCLLDAPGAA 433  
Db 361 PGRNSCPSP--GAPAKTCIMEASTDPLPLNFSNCSRRALBKALDGMGSCLEFER---- 413  
QY 434 LPLPTGLPGMALYQLDQCRQIFGPDFRHCPTNSADVCAQLMCHTDGAEPLCHTKXGS 493  
Db 414 --LPSLPP--MAAF-----GNNMFVEBEGQD-----CGFL--DDCVDPCC----- 448  
QY 494 LPMADGTPC--GPGHL-C-SEGSCLPEEEVERPKPVVDGGAIPWGPWGECS--RTCGGAVQ 548  
Db 449 ----DSLTCQLRPAQCAADGPGCCQNCQL-RP-----SGWQGRPTRGDDCLPEFCPG--- 495

QY 549 FSHRECK-----DPEFONGGR-YCL-GRAPKY-QSCHTEBSP---PDGKSFREQCEKY 596  
Db 496 -DSQCCEPDVSLGGEFCAGGAGVCMHGRCASTYQCCQSLMGPAQPAAPLCLQTAITRIG 554  
QY 597 NAYNYTMDGNLQWVERKYAGVSPDR-C-KLFCRANGSE-----FKVFEAK-- 642  
Db 555 NAFSCGRNPS-----GSYVSCTPRDACGQLQCOT-GRTOPLGSLRDLIMETIDVNGT 608  
QY 643 -----VIDGTLGPEITACVRCQVKAQCDHVVDSPRLDKGV 682  
Db 609 ELNCSWHLIDGSDVAQPLTLTETACGPG--LVCIDHRCQVLDLGAQCRSKCHGHV 666  
QY 683 C 683  
Db 667 C 667

Search completed: May 7, 2004, 11:51.41  
Job time : 30 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:46:40 ; Search time 51 Seconds  
(without alignments)  
5506.102 Million cell updates/sec

Title: US-09-989-687-4  
Perfect score: 4853  
Sequence: 1 MFPAAPAPRWLPFLLLLL.....CNKALKPEDAKPCESQLCPL 890

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_25:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mmc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rhodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_virus:\*  
17: sp\_bacteriophage:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the distribution being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2320.5	47.8	967	4 Q8NE26	Q8NE26 homo sapien
2	2112.5	43.5	759	6 Q8H2M8	Q8H2M8 equus caball
3	1813	37.4	833	11 Q8K384	Q8K384 mus musculus
4	1813	37.4	845	11 Q8BNJ2	Q8BNJ2 mus musculus
5	1794.5	37.0	839	6 Q7Y895	Q7Y895 bos taurus
6	1483.5	30.6	623	11 Q8BGP4	Q8BGP4 mus musculus
7	1475	30.4	2165	5 Q19791	Q19791 caenorhabditis
8	1198	24.7	1688	5 Q8SXB0	Q8SXB0 drosophila
9	1182	24.4	1059	5 Q9WA93	Q9WA93 drosophila
10	1161.5	23.9	1009	11 Q8BKX1	Q8BKX1 mus musculus
11	1161.5	23.9	1600	11 Q8C1B3	Q8C1B3 mus musculus
12	1125.5	23.2	1070	11 Q8CG28	Q8CG28 mus musculus
13	1067.5	22.0	1092	11 Q8BKX1	Q8BKX1 mus musculus
14	990	20.4	900	11 Q8K206	Q8K206 mus musculus
15	976	20.1	769	5 Q8KRL5	Q8KRL5 drosophila
16	836	17.2	183	11 Q9J9P1	Q9J9P1 mus musculus

17	830	17.1	1427	4 Q96137	Q96137 homo sapien
18	740.5	15.3	269	6 Q9GL54	Q9GL54 oryctolagus
19	683.5	14.1	340	11 Q91556	Q91556 mus musculus
20	614.5	12.7	1280	11 Q9EPX2	Q9EPX2 mus musculus
21	592.5	12.2	203	11 Q8CH80	Q8CH80 cavia porce
22	579.5	11.9	1235	4 Q95428	Q95428 homo sapien
23	571	11.8	1572	5 Q44938	Q44938 haemophilus
24	562	11.6	1487	5 Q8MPV5	Q8MPV5 caenorhabditis
25	562	11.6	1558	5 Q81710	Q81710 caenorhabditis
26	562	11.6	2167	5 Q76840	Q76840 caenorhabditis
27	547	11.3	3198	5 Q9U8G8	Q9U8G8 manduca sex
28	541	11.1	192	6 Q95N24	Q95N24 equus caball
29	541	11.1	2174	5 Q9GQRO	Q9GQRO drosophila
30	541	11.1	2776	5 Q869A0	Q869A0 drosophila
31	541	11.1	2898	5 Q86829	Q86829 drosophila
32	536	11.0	2772	5 Q9AV44	Q9AV44 drosophila
33	529	10.9	1461	5 Q8VY48	Q8VY48 caenorhabditis
34	498	10.3	377	11 Q8BL10	Q8BL10 mus musculus
35	495.5	10.2	439	4 Q8N6G6	Q8N6G6 homo sapien
36	493.5	10.2	766	4 P82987	P82987 homo sapien
37	490.5	10.1	1091	5 Q9W1Z6	Q9W1Z6 drosophila
38	488.5	10.1	525	4 Q86RW4	Q86RW4 homo sapien
39	482	9.9	339	4 Q8NEK2	Q8NEK2 homo sapien
40	473.5	9.8	872	5 Q22580	Q22580 caenorhabditis
41	471.5	9.7	497	11 Q8BZD1	Q8BZD1 mus musculus
42	453.5	9.3	1036	11 Q80T21	Q80T21 mus musculus
43	450	9.3	951	4 Q86TH1	Q86TH1 homo sapien
44	446	9.2	187	6 Q95N23	Q95N23 equus caball
45	444.5	9.2	957	11 Q7TSK7	Q7TSK7 mus musculus

## ALIGNMENTS

RESULT 1  
ID Q8NE26 PRELIMINARY; PRT; 967 AA.  
AC Q8NE26;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RA Strusberg R.;  
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.  
DR GO; GO:0004222; F:metallopeptidase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR006586; ADAM\_cysteine.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR InterPro; IPR001590; Peptidase\_M12B.  
DR InterPro; IPR002870; Peptidase\_M12B\_N.  
DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP 1.  
DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
DR Pfam; PF01421; Repolysin; 1.  
DR Pfam; PF00990; tep\_1; 3.  
DR PRINTS; PF01705; TSP1REPEAT.  
DR SMART; SMO0608; ACR; 1.  
DR SMART; SMO0209; TSP1; 3.  
DR PROSITE; PS02015; ADAM\_MEPPO; 1.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
DR PROSITE; PS0092; TSP1; 3.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.

KW Hypothetical protein.  
SQ SEQUENCE 967 AA, 105387 MW, PFIDJ99674201CD CRC64;

Query Match 47.8%; Score 2320.5; DB 4; Length 967;  
Best Local Similarity 49.6%; Pred. No. 6.7e-185;  
Matches 451; Conservative 146; Mismatches 250; Indels 63; Gaps 22;

QY 11 LPELILLILLPLARGA PAPAAGGASELVV--TLPSSAGEMLHLSAFKGVLR 68  
DB 33 VPTLLILAAALIAVS-DALGRPS--EDBELVPELEAPGH-GTRRLRLHAFQGDLE 88  
QY 69 LAPDUSFLAPEKIERLGGSGRATG---ERLRCGFSGTYNGEPESIAVLCRL 123  
DB 89 LRPDSSFLAPGFTLQNV---GRKSGETPLPETDLAHCFYGTGVNGDSSAALSLCEGV 145  
QY 124 SSGFLDGEFTIQOGAGSLAOPHRLQRMGPAPARLP-----RGPWEV---ET 172  
DB 146 RCAFYLLGSAVFIQPLPAAS---ERLATAAPEKPAPIQFHLRRNRQDVGTCGV 200  
QY 173 GEGQROERGDHGEDESEESOEBAEAGASEPP-----PPLGATS-RTKRFVSEARFV 222  
DB 201 VDDEPRPTGKATEDDEDETEDEGEDEGPQMSPODPALQGVGPTGSGIRKGRFVSHRYV 260  
QY 223 ETLIVADSMMAFYGADIQNHILITMSVAARIYKSPISKINIMVYKVLIVDEKMGPE 282  
DB 261 ETLIVADSMMAFHSGLKHYLLTFVVAARLYKSPISIRNSVSLVYKVLIVHDEKMGPE 320  
QY 283 VSDNGGLTLNFCNNQREFNCPSDRPHPHYDTAILLTRONFCQGBGLCDTLGVADIGTC 342  
DB 321 VTSNAILITRNFCNNQREFNCPSDRPHPHYDTAILLTRONFCQGBGLCDTLGVADIGTC 379  
QY 343 DPNSGCSYIEDEGLQAAHTLAHELGHVLSMHPDSEKPTCTLPFGMGKHHVAPLFHLNQ 402  
DB 380 DPNSGCSYIEDEGLQAAHTLAHELGHVLSMHPDSEKPTCTLPFGMGKHHVAPLFHLNQ 439  
QY 403 TLPMSGCSYIEDEGLQAAHTLAHELGHVLSMHPDSEKPTCTLPFGMGKHHVAPLFHLNQ 462  
DB 440 SGPMSGCSYIEDEGLQAAHTLAHELGHVLSMHPDSEKPTCTLPFGMGKHHVAPLFHLNQ 497  
QY 463 HCPNTSAODVCAQLWC-HTGAEPICHTKNGSLPMADGTPCGPGLCSGSGCLDEEVEER 521  
DB 498 HCPD--AASTCSTLWCTGTSGGLVLCQTKH--FPMADGTPSGEGKMCINGKCVKTRKH 553  
QY 522 KRPVVDGMAFWPMGEGSRTCGGVOPSHRECDPEBNGARYCLGRARYOSCHTEEC 581  
DB 554 PTPPHSGVMGMPGDBSRTCGGVQVITRECDNPEKNGSKIEGKRVYRSCNLEDC 613  
QY 582 PP-DGKSPREOQCEKXNAVNTDM-DGNLLQWVPKAGVSPDRCKLFCRARGSEFKVF 639  
DB 614 PNNNGTFREBQCEAHNEFSKASFGSGRAVEMIPKAGVSPDRCKLCOAKGIGYFVL 673  
QY 640 EAKVIDGTICGPELTALCYRGCCVCAACDHVDSRKLDKGVCGGCKNSCRKISGILTP 699  
DB 674 QPKVVDGTPCSTDSVCAQGVACDRIIDSXKFDKCGVCGGNSSTCKKISGIVTS 733  
QY 700 TNYGVNDIYTPAGATNIDVKSHPGVNDGNATYALKTADGOYLGNLISAIEODIL 759  
DB 734 ARPGHNDITITETGATNIEVKORNGSRNNGSFLAKADGTIYLANDYTLSTLEODIM 793  
QY 760 VGTILTKSGSIATIERLOSFRPLPEPLTVOLLTVGAEVFPKVKYTFEVDNDVFSNOS 819  
DB 794 YKGVAVLYRSGSSAALERIRSFPLKEPLITIVLTV-GNALRPKIKYTYFV----- 842  
QY 820 SKERATNTIOPLLHAWYLIDMSSECSSTCGAGMORRTYVEGCDPSGQASATCNALXED 879  
DB 843 KKKKESFNAL-PTPSA-WIIEEMGECSSKCELGMRRLVECRDINGQASBCEAEVFRAS 900  
QY 880 AKPCSQICP 889  
DB 901 TRPCADHPCP 910

RESULT 2

QSHZMS PRELIMINARY; PRT; 759 AA.

ID QSHZMS  
AC QSHZMS  
DT 01-MAR-2003 (Tremblrel). 23, Created  
DT 01-MAR-2003 (Tremblrel). 23, Last sequence update  
DT 01-OCT-2003 (Tremblrel). 25, Last annotation update  
DE Metalloprotease (Fragment).  
GN ADAMS-1.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Beorhume D., Russell D., Richards J.S., Stirois J.,  
RT "Coordinated Regulation of Transcripts Encoding ADAMS-1 and  
RT Progesterone Receptor in Equine Preovulatory Follicles,"  
RL Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF541975; AANI7331.1; -  
DR GO; GO:0004222; Fmetalloendopeptidase activity; IEA.  
DR GO; GO:0008270; Zinc ion binding; IEA.  
DR GO; GO:0006508; Proteolysis and peptidolysis; IEA.  
DR InterPro; IPR006586; ADAM\_cysteine\_M12B.  
DR InterPro; IPR001590; Peptidase\_M12B.  
DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP1.  
DR Pfam; PF01421; Reprolysin; 1.  
DR Pfam; PF00090; TSP1; 3.  
DR PRINTS; PR01705; TSP1REPEAT.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00209; TSP1; 3.  
DR PROSITE; PS50215; ADAM\_MPRO; 1.  
DR PROSITE; PS50092; TSP1; 3.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
FT NON TER 1  
SQ SEQUENCE 759 AA; 83410 MW; C1B84C048918CB9 CRC64;

Query Match 43.5%; Score 2112.5; DB 6; Length 759;  
Best Local Similarity 52.6%; Pred. No. 1.2e-167;  
Matches 380; Conservative 131; Mismatches 174; Indels 37; Gaps 13;

QY 185 EDESEESOEBAEAGAS-----BPP-----PLGATS-RTKRFVSEARFVETLLVADA 230  
DB 1 DDETPSRGAEPRPGQAGTOMAPRDRAFORAGRPCTGSIKRXRFVSSRYVETMLVADO 60  
QY 231 SMAAFYGADIQNHILITMSVAARIYKSPISKINIMVYKVLIVDEKMGPEVSDNGIL 290  
DB 61 SMAEPGSGIKHYLLTFVVAARLYKSPISIRNSVSLVYKVLIVYEBQGPVTSNAULT 120  
QY 291 LRNFCNNQREFNCPSDRPHPHYDTAILLTRONFCQGBGLCDTLGVADIGTICDPNSCSV 350  
DB 121 LRNFCNNQREFNCPSDRPHPHYDTAILLTRONFCQGBGLCDTLGVADIGTICDPNSCSV 179  
QY 351 IEDEGLQAAHTLAHELGHVLSMHPDSEKPTCTLPFGMGKHHVAPLFHLNOTLPMSPCS 410  
DB 180 IEDEGLQAAFTTAHELGHVFNPHDPAKQACASINGVNRDSHMAASVLSLDSQWSPSCS 239  
QY 411 AMTLTLLDGHDCCLLDAPGAALPLPTGLPGRMLYQLDQOCROIFGDPFHCPTSAQ 470  
DB 240 AYMTSLFLDNGHSECLMDRQSTQLPSDLPG--TLYDNRQOQFTFGESKCPD--AA 295  
QY 471 DVCAQLWC-HTGAEPICHTKNGSLPMADGTPCGPGLCSGSGCLDEEVEERPKVVDG 529  
DB 296 STCTTLMCTGTSGLLVLCQTKH--FPMADGTPSGEGKMCINGKCVKTRKHFDTPVHGS 353  
QY 530 WAPMGMGCSRTCGGVOPSHRECDPEBNGARYCLGRARYOSCHTEECPP-DGKGF 588  
DB 354 WAPMGMGCSRTCGGVQVITRECDNPEKNGSKIEGKRVYRSCNLEDCPDNNGKTF 413  
QY 589 REOQCEKXNAVNTDM-DGNLLQWVPKAGVSPDRCKLFCRARGSEFKVF EAKVIDGT 647  
DB 414 REOQCEAHNEFSKASFGSGRAVEMIPKAGVSPDRCKLCOAKGIGYFVLQPKVVDGT 473

```

QY 648 LCGEETLACVAGCQVACGCDHVVDPKLDKCGVCGKGNCKKVSGLTPNYGNDI 707
D 474 PCSPDSISVCCVCGCQVACGCRITIDSKKPKKCGICGANGSTCKKISGTSKAPGNDI 533
QY 708 VTIPAGATNIDVQRSHPGVQNDGNYLALKTADQOYLINGNLASIAIEDILVKGTILKY 767
D 534 VTIPGATNIEVKQRNQRNGSRNNGSFLAKADGTIILNGDFTLSTLEQDITVKGSVLR 593
QY 768 SGSTATLERLQSPFLPEPLTVOLLTVGGEVFPKVKYTFPVDPVDFSMQSSKERATN 827
D 594 SGSSAALERIRSFSPLEPLTVOLLTV-GNALRKIKITTYFV-----KKKESFN 642
QY 828 IIPBLHAQWYLGDMSECSSTCGAGWQRRTVECRDPGQASATCNKALKEPDAKPCESOL 887
D 643 AIPF--SEWVIEEWGECSSKCGQWQRRLVECRDINGOPASBCKAEKVPASTRPCADLP 700
QY 888 CP 889
D 701 CP 702

RESULT 3
ID Q8K384 PRELIMINARY; PRT; 833 AA.
AC Q8K384:
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Similar to a disintegrin-like and metalloprotease (repolysin type)
DE with thrombospondin type 1 motif, 4.
GN ADAMTS4..
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Straube R.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027773; AAH27773.1; -.
DR MGD; MGI:133949; Adamts4.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_Bs.
DR InterPro; IPR008084; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp_1; 1..1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS02015; ADAM_MERO; 1.
DR PROSITE; PS00992; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR Integrin; Protease; Metalloprotease.
KM Integrin; Protease; Metalloprotease.
SQ SEQUENCE 833 AA; 90097 MW; 4CEC83DFEC3AA619 CRC64;

Query Match 37.4%; Score 1813; DB 11; Length 833;
Best Local Similarity 45.0%; Pred. No. 1.6e-142;
Matches 370; Conservative 129; Mismatches 247; Indels 76; Gaps 23;

```

```

QY 126 SFLDDEEFTIOP-----QGAGSLAOPHRLOKRWGPARGARLPFGPEMEVETGEGOROE 179
D 148 VQGYGAELHLOPLLGALNSAGPGA--HILRRSPASS---OGPMCTVAPBSGS--- 198
QY 180 RGDHODESEESQEEBAGASEPPLIGATSTRTKFSVSEARFETLLVADASMAAPYGAD 239
D 199 -----PSPI--SRRTKRFASISREVEITLVVADDGAAFHGIG 233
QY 240 LQNHILLMSVARIYKHPISINSLMWVVKLYIEDSKMGVEVNDGGLTLRNCCNOR 299
D 234 LKRYLTVMAAARAKFKHPSIRNPVNLVYTRLVILGSGEGEQVPSAAQTLRSCTWOR 293
QY 300 RENOPDRPEHYDPAIILLTRQNFQGOEGLCDLIGVADIGTICDPNKSQSVIEDGELOAA 359
D 294 GLNTPNDSDPDHPTAILFTRODLG-VSTCDTLGMAVGTCDPARSCAIVEDGLQSA 352
QY 360 HTLAHELGHVLSMPDDSKPCTRLFGPMG-KHHNMAPLVHNLQTLPSMPCSAAMLTELL 418
D 353 FTAHHLGHVFMNLHDNSKPCNTINGQGSRRHVA PMAHVDPPEPWSPCSAFTITPL 412
QY 419 DGAHGDCLDAPGAALPLPTGLPGMALYOLDQOCROIFGDPFRHCPTNSAQDVCAOLMC 478
D 413 DNGYGRCLLDKPEALHLPATFPGKD--YDADRQQLTFGPDSSHPOLPP--CALMC 468
QY 479 --HTDGAEPYCHTKNGSLPMADGTCGPRHLCSBSCLEPEEVEPKVVDGALPMGFW 536
D 469 SCHLNG-HAMCQTKHS--PMADGTCGSSQA CMGRCCLHVDLKD FNPVQAQGMGPWGW 525
QY 537 GCSRTCGGQVQFSRECKDPEPONGRCYGLGARAYKOSCHTEPC-PDGKSFREQCEK 595
D 526 GDCSRICGGGVQFSSRDCRPVPRNGKCYCBERRRFRFSCTNENPHSALTFREEQCA 585
QY 596 YNAVNTMDGNTL--LQWPKIAGVSPDRCKLFCRARGSEFVFPAKYVDITGLCEB 652
D 586 YN--HRTDLFKSPFPMVMVPRYTVAVAPRDOCKLTCARALGYVYVLEPRVADGTPCSPD 643
QY 653 TLACVAGCQVACGCDHVVDPKLDKCGVCGKGNCKKVSGLTPNYGNDIVTTPA 712
D 644 TSSVCVQGRCHVAGCDRIIGSKKPKKCGICGANGSTCKKISGTSKAPGNDI 703
QY 713 GATNIDVQRSHPGVQNDGNYLALKTADQOYLINGNLASIAIEDILVKGTI-LKYSSTI 771
D 704 GATHILVRQGGSGGLKS--IYALATLSDGSVALNGEYILMSPSTVLPVCAVLSRYSAT 761
QY 772 ATLRLQSPRLPEPLTVOLLTVGGEVFPKVKYTFPVDPVDFSMQSSKERATN 827
D 762 ASETLSGHGFLAQPLTVOLLTVAGNPONARLRYSFVPRPV 802

RESULT 4
ID Q8BNJ2 PRELIMINARY; PRT; 845 AA.
AC Q8BNJ2:
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease.
GN ADAMTS4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Body;
RC MEDLINE=22354683; PubMed=12466851;
RA The RANKO Consortium;
RA The RANKO Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
EMBL; AK083534; BAC38944.1; -.

```

DR MGD, MGI:1339949; Admets4.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR006586; ADAM cysteine.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01421; Repolysin; 1.  
 DR Pfam; PF00090; tsp 1; 1.  
 DR PRINTS; PRO1705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE; PS50215; ADAM\_PROPRO; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR SEQUENCE 845 AA; 91239 MW; B496C3190D1A9225 CRC64;

Query Match 37.4%; Score 1813; DB 11; Length 845;  
 Best Local Similarity 45.0%; Pred. No. 1.7e-142;  
 Matches 370; Conservative 129; Mismatches 247; Indels 76; Gaps 23;

15 LLLLLLLPLARGAPRPAAGQASELVVPTRLPGSA-----GELALHLSAPEGKGV 66  
 46 LLLLLLFLSLA--WPASFLP--REBEIVFPEKLNSSILPGSGVAPRLVRLPAFGML 101  
 67 LRLAPDSFLAPERFKIERLGGSGRATGGERGLRGCFEGITVNGEPESLAVALCRG--LSG 125  
 102 LELEDPGVQVEGLTVQYLQAPLWLGAE--PGYILGTGNGPEBSVASLHMDGALLG 159  
 126 SFLDGESEFTIOP-----QGAGSLAQPHLQRMWPAAPRLPRGPEWEVTEGQROE 179  
 160 VLYRGALHLQPLEGALNSAGPGA--HILRRKSPAS--QGEMCTYKASGS--- 210  
 180 RGDHODEBESEGEAEAGASEPPPLGATSRTRKFSSEARFVETLLVADASMAAFGAD 239  
 211 -----PSPI--SRTRKFSLSRFVETLLVADAKMAAFRG 245  
 240 LQNLITLMSVARIYKPSIKNSINLMVVKYLVEDEKMGPEVSDNGSLTRNFCMOR 299  
 246 LKRYLLTYMAAAAKAFKPSIRNPVNLVTRVLVILGSGQEGPQYGPBAQTLRSFCTWQR 305  
 300 RPNQSDHPEHYTAIILLTRQNFQGEGLCDITGVADIGTICDPNKSCTYIEDEGLQAA 359  
 306 GLNTPNDSDPHFDIAILFTRODLG--VSTCDITGLMDVGVTCDPASCAIVEDDGLOSA 364  
 360 HTLAHELGHVLSMPHDSKPCSTRLEGPWG--KHNYMAPLFVHLNQTLPSPCSAMYTELL 418  
 365 FTAAHELGHVFMHLDNKSCTNLNAGQSSRAHWAPMAHVDEEPMSCSAFTIDFL 424  
 419 DGHGDCLLDAPGALPLPTGLPGEMALYQDDQCRQIFGDFHNCNTSAQDVCAQWMC 478  
 425 DNGYHCHCLLDPEAPLHLPTATPGKD--YDADROCOLTFGSDSHCPQLPPP--CALIMC 480  
 479 --HTDGAEPILCHTKNGSLPMAADGTPCGPGHLCSEGSCLPEBEVARKPVVDGGAAPGW 536  
 481 SGHLNG--HAMCQTKHS--PMADGTPCGSSQACMGARCLHVDQLQDFNVPAQGGGPGPW 537  
 537 GECSTCGGVOFSHRECKDEPONGRGYCLGRBAKYOSCTHEBCP--PDGSAFEEQCEK 595  
 538 GDCSTCGGVOFSRDCTRPVPRNGKCYCEGRTRFRSCNTENCCHPGSALTFREBCAA 597  
 536 YNANYTMDGNL--LOWPKYAGVSRDRCKPCARAGSEKVEAKYIDGTLGPE 652  
 538 YN--HRTLFSFPGPMWVPRYTGVARDDCKLTQARALGYVYVLEPRVADGTPGSPD 655  
 653 TLATCVRGQCVKAGCDHYVDSPRKLDKGVGCGKNSGCRKVSGLTPTNGYNDIVIPA 712  
 656 TSSVVGRCCHAGCDRIIGSKKFKDKCMVCGGSGSCSKSGSFKKFRGYSVVTIIPA 715  
 713 GATNIDVKORSHPGVQNDGNTLALKTADGQYLLNGNLAIISALIEDIIVKGTI-LKISGSI 771

Db 716 GATHILVROQGGSLKNS--IYALKLKSDSVALNBEYTLMPSPDVLPGAVSLRYSGAT 773  
 QY 772 ATLERLOSFPPLPEPLTVQLTVPGVEFPFKYTFEVPNDV 813  
 Db 774 ASETLSHGFLAQLPITLQVL--VAGNPQARLRLYSFFVPRV 814

## RESULT 5

QY95 PRELIMINARY; PRT; 839 AA.  
 AC QY95; 839 AA.  
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Aggrecanase-1.  
 GN ADAMTS-4.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID:9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Arai M., Anderson D., Annis B., Collins-Racie L., Corcoran C.,  
 RA Di Blasio-Smith E., Morris E., Dörner A., Lavallie E.,  
 RT "Cloning and characterization of bovine aggrecanase-1."  
 DR Submitted (May-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL:AF516915; AAF47196.1; --  
 SQ SEQUENCE 839 AA; 90279 MW; 1B488A27DF5B96B1 CRC64;

Query Match 37.0%; Score 1794.5; DB 6; Length 839;  
 Best Local Similarity 45.0%; Pred. No. 5.8e-141;  
 Matches 367; Conservative 124; Mismatches 254; Indels 71; Gaps 22;

15 LLLLLLLPLARGAPRPAAGQASELVVPTRLPGSA-----GELALHLSAPEGKGV 67  
 39 LLLLLLPLPSA--WPASFLP--REBEIVFPEKLNSSILPGGAPARLLVRLPAFGML 94  
 68 RLAPDSFLAPERFKIERLGGSGRATGGERGLRGCFEGITVNGEPESLAVALCRG--LSGS 126  
 95 ELEKPGVQVEGLTVQYLGRAPELLGAE--PGYILGTGINDPEBSVASLHMDGALLGV 152  
 127 FLIDGESEFTIOP--QGAGSLAQP--HILQRMWPAAPRLPRGPEWEVTEGQROE 182  
 153 LQYRGTEHLIOPLEGAEPNSAGPAHLRRKSPSG--QGEMCTYKASGS--- 197  
 183 HQDESEGESEAEAGASEPPPLGATSRTRKFSSEARFVETLLVADASMAAFGADLQ 242  
 198 -----APRGRPSPPRAKRFASLSRFVETLLVADAKMAAFHGAQLR 240  
 243 LITLMSVARIYKPSIKNSINLMVVKYLVEDEKMGPEVSDNGSLTRNFCMORFN 302  
 241 YLLTMAAAAKAFKPSIRNPVNLVTRVLVILGSGQEGPQYGPBAQTLRSFCTWQR 300  
 303 QPSDHPREHYTAIILLTRQNFQGEGLCDITGVADIGTICDPNKSCTYIEDEGLQAA 362  
 301 TPDDADPGHFDIAILFTRODLG--VSTCDITGLMDVGVTCDPASCAIVEDDGLOSA 359  
 363 AHELGHVLSMPHDSKPCSTRLEGPWG--KHNYMAPLFVHLNQTLPSPCSAMYTELL 421  
 360 AHELGHVFMHLDNKSCTNLNAGQSSRAHWAPMAHVDEEPMSCSAFTIDFL 419  
 422 HGDCLLDAPGALPLPTGLPGEMALYQDDQCRQIFGDFHNCNTSAQDVCAQWMC--H 479  
 420 FGHCLLDPEAPLHLPTATPGKD--YDADROCOLTFGSDSHCPQLPPP--CALIMC 475  
 480 TDGAEPILCHTKNGSLPMAADGTPCGPGHLCSEGSCLPEBEVARKPVVDGGAAPGW 539  
 476 LMG--HAMCQTKHS--PMADGTPCGPAQACMGARCLHVDQLQDFNVPAQGGGPGW 532  
 540 SRTCGGVOFSHRECKDEPONGRGYCLGRBAKYOSCTHEBCP--PDGSAFEEQCEK 598  
 533 SRSCTCGGVOFSRDCTRPVPRNGKCYCEGRTRFRSCNTDCTPGSALTFRBEOCAAYN- 591



```

QY 589 YNTYMDGNL---LQWPKYAGVSPRCKLFCARGRSEKFEAKYIDGTLGPEPTLA 655
Db 592 -HRDLEFNFGPMQWVRYGVAPRDCKLTCQTRALGYVLDPRVADGTPCSPSS 650
QY 656 ICVAGQCYKAGCDHVVDPRKLDKGVCGGKSGKRSKVSGLPTNNGYNDIYTPAGAT 715
Db 651 VCVCGRCTHAACDRITGSKKFKDKCMVCGDGSSCSKSGSKFRGYNVVTIPAGAT 710
QY 716 NIDVKQSHPGVQNDGNLALKTADGQYLNGNLALSAIEDLIWKGTI-LKYSGLATL 774
Db 711 HILVRCQGSFVSRS--LYIALKLPGSYALNGEYTLIPSPDVLPQAVSLRYSGLATAAS 768
QY 775 ERLQSFRLPERLTVOULTVPGVPPKVKYTFVFP 810
Db 769 ETLGSGHPLAEPLTLQVL-VAGNPNMARLRYSPFVP 803

RESULT 6
Q8BGP4 PRELIMINARY: PRT, 623 AA.
AC Q8BGP4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease (Fragment).
GN ADAMTS5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Cerebellum, and Retina;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RT Nature 420:563-573 (2002).
RL EMBL; AK084474; BAC38503.1;
DR EMBL; AK084474; BAC38503.1;
DR MGI; MGI:1346321; Adamts5.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008318; F:protein prenyltransferase activity; IEA.
DR GO; GO:0018346; P:protein amino acid prenylation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR002088; PPTA.
DR InterPro; IPR008884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; Lep_1; 2.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SMO0209; ACR; 1.
DR SMART; SMO0608; ACR; 1.
DR PROSITE; PS50215; ADAM_MPRO; 1.
DR PROSITE; PS00904; PPTA; 1.
DR PROSITE; PS50092; TSP1; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER
SQ SEQUENCE 623 AA; 68681 MW; 80CE9BCDAC37502 CRC64;

Query Match 30.6%; Score 1483.5; DB 11; Length 623;
Best Local Similarity 44.7%; Pred. No. 4e-115;
Matches 284; Conservative 93; Mismatches 238; Indels 21; Gaps 10;
QY 261 KNSINLWYKVLIVDEKMGVSDNGGLTIRNFCMQWRFNQSPDRHPEHYTALLTR 320
Db 1 ENHRLAVKVVVLTDRDTSLEVSNNATTLKNCCKQOHQNHQDGDDEHYDAALIFTR 60

```

```

QY 321 QNFCGEGLCDTLGVADIGTICDPNKSQSVIED3GLQAHTLAHELGHVLSMPEHDSKFC 380
Db 61 EDLCGHHS-CDTLGMADVITCISPERSCAVIEDDGHAAFTVAHEIGHLIGLSDHDSKFC 119
QY 381 TLTFGPMGHNHMAPLFTALNLTLMSPSCANVLTLELDGGGDDCLLPAGAAPLPLETL 440
Db 120 EENFGTTEKRLMSLTLSDASKMSKTSATTFEFLDDGKGLDLPRQIIIGPEEL 179
QY 441 PGRMALYOLDQOCROIIFGDFRHCPTSAQDYCAQLMCH-TDGAELCHTKXGSLPMADG 499
Db 180 PGQ--TYDATQCCNLTFEGEYVCP---GMVDCARLMCAVVRQGMVCLTK--KLPAVAG 232
QY 500 TPQPGPHLCSSESCLPEEVEERPKVVOGMAPMPCWGCSTCGGVQFHSRECKDPER 559
Db 233 TPQGRVCLQGHCDVTKTKKTYSTSHGNMGSWGMGQCSRSQGGVQFAYRHCNNPAP 292
QY 560 QNGRYCLGRAPKAYQSCHTTECPDQKSPREOCCEKYNANYTMDG-NLQWPKYAG 617
Db 293 RNSGRCTGRALVYNSCSVTPCPBGKSPRHQCAKNGYQ-SDAKGYTFVEMVPKYAG 351
QY 618 VSPDRCKLFCARGRSEKFEAKYIDGTLGPEPTLAICVAGCYKAGCDHVVDPRKLD 677
Db 352 VLPADCKLTCRAKGYVYVSPKVTDTGTCRPSNSVCVARGCVRGCDGIIIGSKLY 411
QY 678 DKGVCGGKSGKRSKVSGLPTNNGYNDIYTPAGATNIDVKQSHPGVQNDGNVLAALK 737
Db 412 DKGVCGGKSGKRSKVSGLPTNNGYNDIYTPAGATNIDVKQSHPGVQNDGNVLAALK 471
QY 738 TADGQYLNGNLALSAIEDLIWKGTILKYSGLATLEL--QSPRLPEPLTVQULTVP 795
Db 472 KKTGEYLINGKYMISFETIIDINGTVNYSQMSHSDPFIHQMGYSATREILIVQLAND 531
QY 796 GEVFPK--VYTFPVENDVDVFSMQSKERATVNIQILHAQVLTGPMSCSSSTCGAG 852
Db 532 ---PTKALDVRYSPFVPKTKTKQKNSVSHQSNKVGPHSTLQWTVGPMVLACSRCDTG 587
QY 853 MQRRTVECDPSGQASATCNKALPEDAKPCESQJC 888
Db 588 MTRRTVQCQDGNRKLAKGLLSQPSAFKQCLIKKC 623

RESULT 7
Q19791 PRELIMINARY: PRT, 2165 AA.
ID Q19791; Q27524;
AC Q19791; Q27524;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE F25H8.3 Protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berke M., Boulfield J., Burton J., Connell M., Copesey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R., Smalton N., Smith A., Sonhammer E., Staden R., Sulston J., Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.,
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

```

RT elegans. ;  
 RL Nature 368:32-38(1994).  
 DR EMBL; 269361; CAA93288.1; .  
 DR EMBL; 269360; CAA93288.1; JOINED.  
 DR EMBL; 269360; CAA93287.1; .  
 DR EMBL; 269361; CAA93287.1; JOINED.  
 DR F1R; 121371; 121371.  
 DR HSP; P15167; 1DTH.  
 DR MEROPS; M12.135; .  
 DR WormRep; F25H8.3; CE05729.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:000508; F:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR000884; TSP1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; tsp 1; 14.  
 DR SMART; SM00209; TSP1; 18.  
 DR PROSITE; PS50215; ADAM\_MEROPS; 1.  
 DR PROSITE; PS50092; TSP1; 15.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR SEQUENCE 2165 AA; 244397 MW; FCC3DABA9CA888 CRC64;

Query Match 30.4%; Score 1476; DB 5; Length 2165;  
 Best Local Similarity 34.1%; Pred No. 1,1e-113;  
 Matches 327; Conservative 150; Mismatches 319; Indels 164; Gaps 29;

46 RLPGSAGELALHLAFAKGKGVLRALPDDSL-----APEFKIERLQGS 88  
 DB RLPQVARDP-----GHACHLRILSDDAVYVHLRWNQIPDSHNKSVPHFSNPAFM 148  
 QY 96 RLPQVARDP-----GHACHLRILSDDAVYVHLRWNQIPDSHNKSVPHFSNPAFM 148  
 QY 89 GRATGGERGLRG-----CFPGSTNAGEPSLAVALC---RGLSGFLDDEHFFIQ 137  
 DB 149 VLYLDESEVERGKMSRTDPDCTYRAHVGVQH-SIVNLCESEGLYMLALPGSHIVE 207  
 QY 138 P-----QAGAGSLAOPHRLQWGPAGARPLRGPENEVEGEGQERGDH--QEDSEE 189  
 DB 208 PLSNGTNEHDGASHRHQHLVRKPKPFKSP-----DHLNSTSVN 248  
 QY 190 EECSEBACASPPPLGATSTKFPV-SEAFVETTLVAASAAPGADLQNHILTM 248  
 DB 249 EETETVAWQDQMEDVIERKASRAASMDHYEVLVVADTKYEVYGRSLSELYVTLF 308  
 QY 249 STAAAIYKAPSIKNSINLMVYVLLVEDEKWPESVDNGTLTLNFCWQRRFPQPSDRH 308  
 DB 309 STVAATYRHQSIRASINNVVLYLKTENAGFRITQNAOQLDQFCWQYVNDPDDSS 368  
 QY 309 PEHYDTAILTRQNTFGQEGELCDLVGVADIGTICDPNKSQSVIEDEGLQAHTLAHELGH 368  
 DB 369 VOHHVVALITFRKDCIRGQKCDTGLAELGTCMCKSCAIIENINGLSAFTTAHEIGH 428  
 QY 369 VLSMHDSPKCTRLRFGMCK-----HHMAPLFYVLTNTLWSPSCSA 411  
 DB 429 VPSIHDERKKS-TYFVNKKCKFQSTKFKDTQONFHMAFLPILNTHPMNSPCSA 487  
 QY 412 MYLTELDDGNG--DCLLDAPCAL-----PLPTGLPGMALYQLDQCRQIFGDFRHP 465  
 DB 488 GWLEFFLENNRGQTCLFDQVERRYEDVFRDEPGK--YDAHQCKFPGASELCP 545  
 QY 466 NTSADVCAQLWCHN--DGAEPICRHQNSLSLWADSTPGSPGH--LGSQSGSL---PEEV 519  
 DB 546 ---YVPTCRMLCATFYSGQNGCRTH--MFWADSTPDEBSRSMFCHHGACVRLAPSLT 600  
 QY 520 EBPKEVVDGVAFWPWCESCRTGCGVQVFSHRECKDEPONGRGYLCGRBAKYQSCHTE 579  
 DB 601 K-----IDGQWDMRSMWGCSPRTGCGVQKGRDSCPKPRNGKRYCVGQERYRSQNTQ 655  
 QY 580 EPPPDGKSRREGCKEKN--ANNYDMNGNLQWPKYAGVSPRRCCLFGRAGRSEFK 637  
 DB 656 EEPWDTQPRREVQCEPNNKDIGIGVASTNTHWPKYANVAPEMRCLYCRLSGSAFY 715  
 QY 638 VFEAKVIGDTLGGPETAICVAGQCVKACDHYVDSPPKLDKCGVCGKSGSKRVGSL 697

DB 716 LNDKTVDDGTPCORNDDICVAGACWPGACDHLHSTRDKCGVCGGDDSCVKYKGF 775  
 QY 698 TPT-NGYNDIVITIPAGANINIDYKQSHPGVQNDGYTLAKPADGYTLNGNATIAIRQ 756  
 DB 776 NEQGTGYEVKMIIPGASNIDIRQGYNNKEDDYLSTLRANGSEFFLNGHFOVSLARQ 835  
 QY 757 DILVKGTLIKYSGSIATLERLOSFRPLPEPLTVQLTVGEVFPKRYTFF---VPNDV 813  
 DB 836 QIAFGDTVALEYSSDILIRINGTPIRSDIVYHLVLSVSH--PDISYEVYTAAPNAV 893  
 QY 814 -----DFSN-QSKERATNITQPL----- 832  
 DB 894 IRPISALYLMRYDTWTEDCRACRQSQSKLMCLDMSHRSQSHRNCCNVIKPKQATRM 953  
 QY 833 ---LHAQWVLGDMSECSSTCGAGWQRTVEGRDPGQ--QASAT---CNKALKPEDAKPC 883  
 DB 954 CNIDCSTRWTEDEVSSCAKCGSGQKRVSCVYKMGDRQTAASEHLCDRNSKPSDIASC 1013

RESULT 8  
 Q8SXB0 PRELIMINARY; PRT; 1688 AA.  
 ID Q8SXB0 Q9VF61;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE GH16393p (Cg14869 protein).  
 GN CG6107.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresner D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guertel H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunco J., Pacled J., Pargass V., Park S.,  
 RA Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celinker S.;  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gedeys J.D.,  
 RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt R.G., Nelson C.R., McKlos G.L.G.,  
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Piankocch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Beyraktaroglu B., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhanderi D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,  
 RA Butts K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Dommes W., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaeser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Moberly C., Morris D., Moshireti A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Massaman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zheng G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*,"  
 Science 287:2185-2195(2000).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RA Ceiniger S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Barzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorselt V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreira S., Frisze E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoekins R.A., Hostin D., Howland T.J.,  
 RA Iobagwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Paclab J., Patagas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
 RT "Sequencing of *Drosophila melanogaster* genome,"  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Maier S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Ceiniger S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frisze E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Seierle S.M., Smith E., Shu S., Smutnak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
 RT "Annotation of *Drosophila melanogaster* genome,"  
 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Ceiniger S.E., Gibbs R.A., Rubin G.M., Venter J.C.,  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL, AY094716; AM11069.1; -.  
 DR EMBL, AE003709; AAF53199.2; -.  
 DR FLYbase; FBgn0038340; CG6107.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001590; peptidase\_M12B.  
 DR InterPro; IPR002870; peptidase\_M12B\_N.  
 DR InterPro; IPR006025; pept\_M12B\_BS.  
 DR InterPro; IPR000884; TSPI.  
 DR Pfam; PF01562; pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR SMART; SM00209; TSPI; 5.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS50092; TSPI; 3.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 1688 AA; 189867 MW; 48FB8DD4DE0CA4D2 CRC64;

DB 197 PLANSPPATPALTFRRLDNDTSTHPPDDGLDDEBHSFVNPFTKYNTSLSEADIT 256  
 QY 47 -----LPSAGEAL-----HLSAFGKGFVRLADDSFL 76  
 DB 257 YEKNSNDINSLTKSASAFATGTFRMNSNEIMDPHQYNLVNFGROHLVLRQDASPV 316  
 QY 77 -----APEKTI-----ERLGSSRAALGGERGLRGCFSGTYNGEPSTLAVALCRLLSGSF 127  
 DB 317 HNSMTHIRILKEGHEHPQETEAEPRL-GCFYSGVEDDPHSXVSLSGDMTGT 375  
 QY 128 LUDCEFTIOPQAGAGSLAEPRLQRMGPAGARP-PPRPEWEVEVTEGEOQROERG-DHOED 186  
 DB 376 KTSFGALLIQPNRISDSDEVLHVRKSGQRARNAV-----SKELGLDDPMS 423  
 QY 187 SEESQEEBAGASAPPPPLGATSTTRFV-----SEAFVETLVADNSMAAFYADLON 242  
 DB 424 KLEQVQEEBOKSKR-----KLNRKKHVADVNDQVYTLLEVLAVDNSMRFHGEDLOP 477  
 QY 243 HILTVSVAARYKPSIKNSINLMVNVKVLVEDEKMGPEVSD--NGGLTLRFNCMQR 299  
 DB 478 YILIMSVISSTFADASIGNSIRILVRLISL-----PNINDQHSNEMLKHFCQF-- 529  
 QY 300 RFPQPSDHPBEHYTALITLTONFCQ--EGLCDTLGVADIGTICDPKSGSVIEDGLQ 357  
 DB 530 -INQGYER---DTAMLTREPIGSGVPKICMVLGLAEIGVCS-SSSGSIVDDTGLP 583  
 QY 358 AANTLAHLGHVLSMSPHDSKPC---RLTGPMKKHVMAPLFVHLNQLTPMSPCSAMY 413  
 DB 584 TAFIWAHEIGHILNNHDDDKCPYTRQNNKYLHMSVWGIIHM-PMWSKCSNHF 642  
 QY 414 LTELDDGHGDCILDAPAPALPLPT-GLPGRMALYOLDQCRQIOPDPFRHCPTNSADV 472  
 DB 643 VSEFLKTKDSKLETSVGAHLYGTERLPGE--IYSLDAQQLSPGDNFGVCP--TDEE 697  
 QY 473 CAQLMCH-TDG-ABELCHTXGSLPMADGTFCG-PGHLCSGSLCPBEVEVRPRVYDGG 529  
 DB 698 CKRLMCKNTSGNSNEQACASN-LPMADGTGSSGHHQCRGKCVSNHGVGRQ--VNGG 753  
 QY 530 WAPMPWMEGCRTCGGVQGFHRECKDPEPNGRGVCLGRBARYCSCTECCPPDGKFSR 589  
 DB 754 WGPWTFPFLPCSLTGGGVQESRRECNQVPENGGKCYCGSKKTKRSCNTHCCPSMDPR 813  
 QY 590 EOCCKYNAVNTYMDGNL-LQWYPKYAGVPRDRCKLFCAGRGSEKVEAKYITDGL 648  
 DB 814 EOCQYAMNGRMNITPGVNPDTKWPKY---EKDACKLFCMDMKVTFMLKSMVTDGTS 869  
 QY 649 GPEPTLACVRGQCVKAGCDHVDSPRKLDKGVCGGKNSCRKVSGLTPTN-YGYND- 706  
 DB 870 CAVDSFKCVNGICRPAQCDNELNSIAKLDKCGVCEBNDTCHEVTGNLVSNLGLDNG 929  
 QY 707 -----IYTPAGATNIDVKORSHPGVNDGNVYALKTADQOYLINGLAISALBOD 757  
 DB 930 NEPNKTLVYVRIPIGASNIITRGYP---DQNFVILTDRDNEMLNGKF-LKTYPLK 984  
 QY 758 ILVKSTILKYSGLATLERLSFR--PLPELTLYOILLVPEVPPK-----KYQTFV 809  
 DB 985 FYVAGVYQVYSSSVQVNTTYSWLSRDLIVQIISL--DVBSRKQDVTLLSYSTTI 1042  
 QY 810 PNDVDF-----SMOSK 821  
 DB 1043 DKRPDYEAELIYVEMQAPNCDLSLCEGRSHRLPACISTTGQVKAPOFCDKSMAPXID 1102  
 QY 822 ERA-----TNIIOPLHLAAQ----- 836  
 DB 1103 DRACNTGRNLVYTSISEGCAAGELGTRKXTYACVQFTTNMQRSNIVMSYCKLAKDV 1162  
 QY 837 -----WVLGWSCESTCGAGWQRTVECRDPSGQAS-ATCNKALPE--DAKP 882  
 DB 1163 AYHEECRGECVVLWSMSTCSKSCGSGQRAHGYLHNSRVSDDLQCNPRTKPHLNTLIGI 1222  
 QY 883 CESQLCP 889  
 | : : |

Query Match 24.7%; Score 1198; DB 5; Length 1688;  
 Best Local Similarity 30.0%; Pred. No. 1,4e-90;  
 Matches 326; Conservative 144; Mismatches 343; Indels 274; Gaps 40;  
 23 PLARGADAPRPA-----GGQASELVVPTP----- 46











RESULT 14			
Q8K206			
ID	Q8K206	PRELIMINARY;	PRT; 900 AA.
AC	Q8K206;		
DT	01-OCT-2002	(TrEMBLrel. 22, Created)	
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)	
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)	
DE	Hypotheetical protein.		
GN	ADAMTS16.		
OS	Mus musculus (Mouse).		
OC	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Straussberg R.;		
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC034739; AA134739.1; -.		
DR	MGD; MGI:2429637; Adamts16.		
DR	GO; GO:0004322; F:metalloendopeptidase activity; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InferPro; IPR001590; Peptidase_M12B_N.		
DR	InferPro; IPR002870; Peptidase_M12B.		
DR	InferPro; IPR000884; TSP1.		
DR	InferPro; IPR008085; TSP1.		
DR	Pfam; PF01562; Pep_M12B_propep; 1.		
DR	Pfam; PF001421; Reprolysin; 1.		
DR	Pfam; PF000690; tsp_1; 1.		
DR	PRINTS; PRO1705; TSP1REPEAT.		
DR	SMART; SMO0209; TSP1; 1.		
DR	PROSITE; PSS0215; ADAM_MPRO; 1.		
DR	PROSITE; PSS0092; TSP1; 1.		
DR	Hypotheetical protein.		
QK	SEQUENCE 900 AA; 100679 MW; 590BE2BE0E73740FF CRC64;		

		Best Local Similarity 31.0%; Pred. No. 1,4e-73;		
		Matches 273; Conservative 135; Mismatches 352; Indels 120; Gaps 35		
QY	26	RGAPAPPAAGGASELVITRLPGSAGELAHLHSAFGKGYLRLAEDDSFLAPERFIETL	85	
Db	84	RAAVTQP--GGATLTL-----RLKGPRIIDLHLDKA-----ASWMAFGWVQTL	126	
QY	86	GGSGRATGSEGLR-----GCFPSGVNDEPESLAVALSLCRLSSGFLDDEEFTIQ	139	
Db	127	G-----KGGTSSVQMFPPPEZNCFYQSILRQGSNSVALSLCCGSLDMITKXTDYFLKR-	180	
QY	140	GAGSLAOPH--FLQR-----WGPAARPLPRGMEVETGEGQOEGRDHQ	185	
Db	181	-----LPHLTSKLNRSAQDPSHYLKRSTERCAPRENEVLTMTKRRLAPHLHND	234	
QY	186	DSE--EESQEEAEAGSE---PRPP-----LATSIXKRFV-----SEARFVETLL	226	
Db	235	NHLLGPSQKQHPGCGRRKRYMQPNDLTYLRLPEYKSSNHRKSLKSRNELINVELTV	294	
QY	227	VADASMAAFYG-ADLQNHILTLMSVAARIYKSPSINKSINLNVVKYLVEDERKGEVSD	285	
Db	295	VYDRKMQSHGHENITTVLTILNMVSALPDQGTIGNNITIVGILLDEDEQGLATISH	354	
QY	286	NGGLTLNPFQWQR--RFQPSDRPHRYDPAIILTLQNPGC--QEGLCDTLGVADITGCD	343	
Db	355	HADHTLTSPCCQOQSLMGKDGTRH---DBAILTLGDIDSMKNBCDITLGAAPISGWS	410	
QY	344	PKMSCSVIEDEGLQAAHTLAHELGHVLSMPHD--DSKPCITLFPQMKHVNAPLFLVHLNQ	402	
Db	411	KXRCTVNEDESLGLAFTLAHESHNRGVHDEGNNCKKSEB-----NINSPTLAHRNG	465	
QY	403	TLPMSPCSAMYTELLDGGHDCILLDPGAA--LPLTGLPGMALYQDQQCRQILFERD	460	
Db	466	VFSWSSCSQYLAKHKLSTAQALICLADDPKVEKYKPEKLPQG--LYDANTQCKMQFGK	523	
QY	461	FRHCEYTSADYQCAQLMCHTDGAERPLCHTKNGSLPMADGTPCCPGHLCEGSLPDEEYE	520	
Db	524	ATLCMLDPKQDI CRYLMCHRIGRK--CETK--TPPAEGLTCCQGDWMCWGQCVAYGD-E	578	
QY	521	RKRPVVDGMAWPGWGECSRTCGGQVQFQSHRECKDPEPONGRYLGRRAKYQSCHTBE	580	
Db	579	GKRP--THGWSWMSWSPSCSRTCCGGTSHSRDLCTNRPBHGKFCQGGSTRLLKCNQR	637	
QY	581	CPDPGSPPEQCEKYNAVNTDMDGKLLQMWPKYAGVSPRDCFLFCARGRSEPKVE	640	
Db	638	CPUDSVDFPAACARYSKRFR--GWLKMKRP-YQLBQBDICLKYCTLAEGDFPFPSL	693	
QY	641	AKYIDGTLCEPBTALCYRQCVYAGCDHYVDSPRKLDKCGVCGGKNSC--RKYSGSL	697	
Db	694	NKYKGTPSEDSMRNVICIDGKCEHVGDDVLTGSDATEDSGVCVKGNNSCVTHRGYSKH	753	
QY	698	TEPNVGRNDYITIPAGATNDIVKRSRSPGVQNDGNTALAKTADGQYLNGNLALISABD	757	
Db	754	HSNIOYHR-MVITPGARSHIYV---TNISTSYISVNSLSKRYLLIGHSVDMPGH-	806	
QY	758	ILVKGITLKYSGISITLERLSQFRPLDEEPLTVQLTPGGEVFPKRYTFPVNDVFSM	817	
Db	807	YKFSGATRYKRSYKEPENLTSPPETMETLIVELF---QGRNPVGAWEFSLP-----	856	
QY	818	QSSKERATNTIIQPLLAQWVLGDMSECSSTCGAGWQRPT	857	
Db	857	RSQAKK--TPAAQP--SYSMAIIV-SECSVSCGGRDRCQF	891	
RESULT 15				
QGMRL5	QGMRL5	PRELIMINARY;	PRF;	769 AA.
AC	QGMRL5;			
DT	01-OCT-2002 (Tremblrel. 22, Created)			
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	GH22104P.			
NN	CC4096.			

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryote, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota;  
 OC Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha;  
 OC Ephydroidea, Drosophilidae, Drosophila.  
 ON NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungai C.J., Nuno J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Pounanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceoliker S.;  
 RI Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, AY119538, XAM50192.1, -;  
 DR FLYbase; FBgn0029791; CG4096.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001590; peptidase\_M12B.  
 DR InterPro; IPR002870; peptidase\_M12B\_N.  
 DR InterPro; IPR006025; pept\_M2n\_BS.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Pep\_M12B\_dropep; 1.  
 DR Pfam; PF01421; Repolysin; 1.  
 DR Pfam; PF00090; tsp\_1; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 769 AA; 86543 MW; 39404FBD271E1968 CRC64;

Query Match 20.1%; Score 976; DB 5; Length 769;  
 Best Local Similarity 31.5%; Pred. No. 1,76-72;

Matches 251; Conservative 103; Mismatches 279; Indels 164; Gaps 26;

QY 9 RWLPLLLLLLLLLLPLARAPAPPAAGQAS-----LVPTRL----- 47  
 DB 11 RWLG--LLHFWLVVVGQGVPRVYGLHSDLVAGGQLVPRRVHPDGAFWTHQLAYA 68  
 QY 48 -----PSAGELALHLSAFGKGFVLRAPDSTFLAPFKTERIGSG 89  
 DB 69 HELDHRHRORRSINSEHDTQAADLHLPLANETLLELMASVFLAPNLVYER----- 123  
 QY 90 RATGGERGLR-----GCFPSGVNGEPESLAVALCRGLSGFLDGEFTI--- 136  
 DB 124 ----HRDRLTRSPPLTRRLNCHFGKVRGQPAITVAISTCAGLVGHIRTAGNEYFIEPS 179  
 QY 137 ---OPQAGG-----SLAQPHRLQWGPAGA-----RPLPRG 165  
 DB 180 KEHEPHVNHPRVVFQRSSVKKPKHSIRKKNKRGKSGSGAEVSNCGTRPEPRRMETR 239  
 QY 166 PEME-----VETGEGQROQRGDHDESESEGEAEAGSEPP-----PPL 206  
 DB 240 LEWQARGKVKVQGGRLRRHHHHHHHHHGHKTRHHQKISRVPTKFKYETQFQTEPDH 299  
 QY 207 GATSRTRFVSEARFVETLLVADASMAAFYADLQNHILTLMSVAARIYKPSIKSINTL 266  
 DB 300 AEIPRRRSISSPRHVETLLVADATMSAFH--RLNGYLLITIMVSAIYKPSIGNSIEI 358  
 QY 267 MMYKVLIVEDEKGRPE--VSDNGGLTLRNCWQRRPNQSDRAHPEHYDTAILTRQNF 324  
 DB 359 VVVRITQLDEESQLQNTLQNAQKLDRCFCSWQHKLNKSGEKDPHHHDVAILLTRKNIC 418  
 QY 325 GQEGLCITLGVADIGTICDPNKSCSVLEDEGLQAHTLAHELGLVLSMPHDSK--PCTRL 383  
 DB 419 ANN--CMTLGLAVGCKRCKXKSCSVNEDNGIMLSHTITHELGNFMGHDTAKIGCHPR 476  
 QY 384 FGPNKGKHVAPLPLFHLNQLTLPWSPCSAMVLTLLDGGHDCLLDAPGALPLP---TG 439

DB 477 VGPI--VAIMPTRGADTLQVCMNSCRKTYTHFLDGLGRCLDDPP---TPLBRYNTG 531  
 QY 440 -LPGEMALYQLDQOCROIFG--PDRHCPTNSADVCAQLMCHTDGAEPICHTKNGSLPW 496  
 DB 532 ELPGWR--YNARGQCRQLFNLITDSEVACGAPHEFGSTLCKVNG--ECVTHMR---PT 584  
 QY 437 ADGTPCGGHLCSGSCCLPEEVEPRFKPVNDGMAPNPMGECGRTGGGVQFQSHRECKD 556  
 DB 585 APGLICGNKWCQNGKCVREEL---AAVNGWGDSEWSESCSRSCGGSVSTQGRBCDN 640  
 QY 557 PEPONGRYCIGRRAKYQSCHEECPPDGKSFREQCEKXNAVYTTDGNILQWVPKXA 616  
 DB 641 FVPANGGVFCIGEKRYKICRKPQPAEPSPFRAGQCARFNVS---GQATYKMLPRFD 697  
 QY 617 GVSPPDRCKLFCRARGSEFVFEAKYIDGTLGPEPTALICVRGQCVKAGCDHYVDSBRK 676  
 DB 698 KNNP---CKLFCSDVDPTIINMGATVLDGTPCTLGTNNMCIDGIC----- 740  
 QY 677 LDKGVCCKGKNSCRKV 693  
 DB 741 --KLVAIGSTRRCRTI 755

Search completed: May 7, 2004, 11:51:07  
 Job time : 60 secs